GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 8, 2004, 09:24:21 ; Search time 129 Seconds
(without alignments)
578.740 Million cell updates/sec

Title: US-10-088-945A-10
Perfect score: 1195
Sequence: 1 MAERDLRLLGAWFSPFALRV.....KLVEFAKILQLKWAAAAAAK 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1155.5	96.7	233	12	US-10-424-599-194349	Sequence 194349,
2	1000.5	83.7	234	12	US-10-424-599-200780	Sequence 200780,
3	624	52.2	233	9	US-09-765-213A-4	Sequence 4, Appli
4	623	52.1	233	9	US-09-765-213A-2	Sequence 2, Appli
5	607	50.8	233	9	US-09-765-213A-6	Sequence 6, Appli
6	573	47.9	233	16	US-10-437-963-153017	Sequence 153017,
7	557.5	46.7	375	16	US-10-437-963-154851	Sequence 154851,
8	542	45.4	306	16	US-10-437-963-131794	Sequence 131794,
9	541	45.3	244	16	US-10-437-963-117929	Sequence 117929,
10	540	45.2	235	16	US-10-437-963-148018	Sequence 148018,
11	535.5	44.8	237	16	US-10-437-963-152885	Sequence 152885,
12	535	44.8	243	16	US-10-437-963-148024	Sequence 148024,
13	534	44.7	236	16	US-10-437-963-132134	Sequence 132134,
14	533	44.6	336	16	US-10-437-963-130681	Sequence 130681,
15	528	44.2	232	16	US-10-767-701-45399	Sequence 45399, A
16	525.5	44.0	233	16	US-10-437-963-109052	Sequence 109052,
17	523.5	43.8	233	16	US-10-776-956-16	Sequence 16, Appl
18	522.5	43.7	237	16	US-10-437-963-204635	Sequence 204635,
19	521	43.6	241	16	US-10-437-963-152637	Sequence 152637,
20	519.5	43.5	233	16	US-10-437-963-186789	Sequence 186789,
21	514.5	43.1	236	9	US-09-765-213A-9	Sequence 9, Appli
22	514	43.0	240	16	US-10-437-963-144283	Sequence 144283,
23	512	42.8	264	16	US-10-437-963-109053	Sequence 109053,
24	511.5	42.8	233	16	US-10-437-963-186791	Sequence 186791,
25	511	42.8	269	12	US-10-425-114-50727	Sequence 50727, A
26	511	42.8	269	12	US-10-425-114-66732	Sequence 66732, A
27	508.5	42.6	229	9	US-09-765-213A-8	Sequence 8, Appli
28	506.5	42.4	354	16	US-10-437-963-151237	Sequence 151237,
29	504	42.2	235	16	US-10-437-963-106610	Sequence 106610,
30	501.5	42.0	241	16	US-10-767-701-45167	Sequence 45167, A
31	501.5	42.0	248	12	US-10-425-114-52797	Sequence 52797, A
32	501	41.9	237	12	US-10-425-114-62288	Sequence 62288, A
33	498.5	41.7	226	16	US-10-767-701-39274	Sequence 39274, A
34	498	41.7	256	12	US-10-425-114-52786	Sequence 52786, A
35	497	41.6	230	16	US-10-437-963-170660	Sequence 170660,
36	493	41.3	240	16	US-10-437-963-202064	Sequence 202064,
37	489	40.9	225	9	US-09-765-213A-7	Sequence 7, Appli
38	486.5	40.7	243	16	US-10-437-963-154488	Sequence 154488,
39	486	40.7	232	16	US-10-767-701-43039	Sequence 43039, A
40	484.5	40.5	245	12	US-10-425-114-72039	Sequence 72039, A
41	483	40.4	245	12	US-10-425-114-53472	Sequence 53472, A
42	483	40.4	245	12	US-10-425-114-72237	Sequence 72237, A
43	480	40.2	234	16	US-10-437-963-161253	Sequence 161253,
44	469	39.2	263	12	US-10-425-114-62056	Sequence 62056, A
45	456	38.2	257	16	US-10-437-963-148917	Sequence 148917,

ALIGNMENTS

RESULT 1
US-10-424-599-194349
; Sequence 194349, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 194349
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(233)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17524C.1.pep
US-10-424-599-194349

CIP of 09/985,678 which is a CON of 09/304,317 (5/6/99) - both parents are CRPE but do not show a hit in pending database so appears Seq. was present only in the CIP

Query Match 96.7%; Score 1155.5; DB 12; Length 233;
Best Local Similarity 97.0%; Pred. No. 7.7e-108;
Matches 226; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

```
Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||:|||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQRDLRLLGAWFSPFVLRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSV-LATE 119
              |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||
Db         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDGANARFWVSYIDDKWLTSLKSVLLAXS 120
```

```
Qy        120 DDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNE 180

Qy        180 RKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAK 232
              ||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        181 RKLLDETKHPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAK 233


RESULT 2
US-10-424-599-200780
; Sequence 200780, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200780
;   LENGTH: 234
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_2332C.1.pep
US-10-424-599-200780

  Query Match             83.7%;  Score 1000.5;  DB 12;  Length 234;
  Best Local Similarity   84.6%;  Pred. No. 3.1e-92;
  Matches  192;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:||||||||||||:|||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||||||||||| |: | |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | |||||||  ||::|||||| ||||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| |||||||||||||:||||:||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 3
US-09-765-213A-4
; Sequence 4, Application US/09765213A
; Patent No. US20020079846A1
; GENERAL INFORMATION:
;  APPLICANT: Facchini, Peter J
;  TITLE OF INVENTION: No. US20020079846A1el Glutathione-S-Transferase Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides and Methods of Use
;  FILE REFERENCE: 22542-001
;  CURRENT APPLICATION NUMBER: US/09/765,213A
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 60/176708
;  PRIOR FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: Papaver somniferum
US-09-765-213A-4

  Query Match             52.2%;  Score 624;  DB 9;  Length 233;
  Best Local Similarity   54.0%;  Pred. No. 2.4e-54;
  Matches  116;  Conservative   39;  Mismatches   60;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::|||||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKIPVMIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|| : ||::|| || :: :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGEEKIGYIDIAFGCYIGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              ||| |||| ||| | ||  || ::|||: |:|||
Db        188 ETKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 4
US-09-765-213A-2
; Sequence 2, Application US/09765213A
; Patent No. US20020079846A1
; GENERAL INFORMATION:
;  APPLICANT: Facchini, Peter J
;  TITLE OF INVENTION: No. US20020079846A1el Glutathione-S-Transferase Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides and Methods of Use
;  FILE REFERENCE: 22542-001
;  CURRENT APPLICATION NUMBER: US/09/765,213A
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 60/176708
;  PRIOR FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: Papaver somniferum
US-09-765-213A-2

  Query Match             52.1%;  Score 623;  DB 9;  Length 233;
  Best Local Similarity   54.0%;  Pred. No. 3.1e-54;
  Matches  116;  Conservative   39;  Mismatches   60;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::||:||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKMPVLIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|||: ||:||| || :: :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGGEKIGYVDIAFGCYVGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              | | |||| ||| | ||  || ::|||: |:|||
Db        188 EEKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 5
US-09-765-213A-6
; Sequence 6, Application US/09765213A
; Patent No. US20020079846A1
; GENERAL INFORMATION:
;  APPLICANT: Facchini, Peter J
;  TITLE OF INVENTION: No. US20020079846A1el Glutathione-S-Transferase Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides and Methods of Use
```

```
;  FILE REFERENCE: 22542-001
;  CURRENT APPLICATION NUMBER: US/09/765,213A
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 60/176708
;  PRIOR FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (58)..(58)
;   OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
;   OTHER INFORMATION: specification
;   NAME/KEY: VARIANT
;   LOCATION: (61)..(61)
;   OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
;   OTHER INFORMATION: specification
;   NAME/KEY: VARIANT
;   LOCATION: (157)..(157)
;   OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
;   OTHER INFORMATION: specification
;   NAME/KEY: VARIANT
;   LOCATION: (163)..(163)
;   OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
;   OTHER INFORMATION: specification
;   NAME/KEY: VARIANT
;   LOCATION: (171)..(171)
;   OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
;   OTHER INFORMATION: specification
;   NAME/KEY: VARIANT
;   LOCATION: (189)..(189)
;   OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
;   OTHER INFORMATION: specification
;   OTHER INFORMATION: Description of Artificial Sequence: Consensus
;   OTHER INFORMATION: Sequence
US-09-765-213A-6

  Query Match             50.8%;  Score 607;  DB 9;  Length 233;
  Best Local Similarity   53.0%;  Pred. No. 1.3e-52;
  Matches  114;  Conservative   37;  Mismatches   64;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::|| ||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKXPVXIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|| : ||: || || :  :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGXEKIGYXDIAFGCYXGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              | | |||| ||| | ||  || ::|||: |:|||
Db        188 EXKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 6
US-10-437-963-153017
; Sequence 153017, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 153017
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_53010C.1.pep
US-10-437-963-153017

  Query Match             47.9%;  Score 573;  DB 16;  Length 233;
  Best Local Similarity   47.6%;  Pred. No. 3.3e-49;
  Matches  107;  Conservative   43;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|:||| | |||||| ::||: ||| || ||| |  |||||| |||||||:||  |  |
Db          7 ELKLLGMWASPFALRAKLALSFKGLSYEYVEEDLKNKSELLLTSNPVHKKVPVLIHNGKP 66

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              :||| :||:|||| : :  : :|| : |||| |||| :||||| | |       : :: |
Db         67 VCESQVIVQYIDEAFPDAGVPLLPSDPYDRAVARFWAAYIDDKLLKSWLQASMCKTEQEK 126

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
                  ::    :  :|  | :||:|| :|||| :|:||:  |: : |:|| | ::  :| |
Db        127 AAAMKETFAAVANLEAAFKECSKGKPFFGGDAVGYVDVTLGAVIGFVRVGEAVHGMRLFD 186

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAA 229
              :: | |  | : |||  | | :||:| :| |:||: | :|||||
Db        187 ASRSPLLDAWLDRFAALDAAKAVLPDTGRLAEYAKMKQAEWAAAA 231


RESULT 7
US-10-437-963-154851
; Sequence 154851, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154851
;   LENGTH: 375
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_54671C.1.pep
US-10-437-963-154851

  Query Match             46.7%;  Score 557.5;  DB 16;  Length 375;
  Best Local Similarity   47.8%;  Pred. No. 2.3e-47;
  Matches  109;  Conservative   46;  Mismatches   62;  Indels   11;  Gaps    4;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              |:||| | ||:  ||::||: ||| || ||| |: |||||| :||||||:||  |  | |
Db        153 LKLLGLWVSPYTHRVKLALSFKGLSYEYVEEDLSNKSELLLSTNPVHKKVPVLIHNGKPI 212

Qy         66 CESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK----WLTSLKSVLATEDD 121
              ||| :||:|:|| : |: :|:|| ::|||| |||| :||:||    || |  |:  ||::
```

```
Db        213 CESQVIVQYLDEEFPNSGVSLLPSDSYDRAIARFWAAYINDKLMPSWLQS--SMGKTEEE 270

Qy        122 EAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERK 181
              |: |  :|  | :  :|  | :||:|| :|||||:|::|:  |: : ::|  | :: |:
Db        271 RAEAL--KQTLEAVANLETAFKECSKGKPFFGGDTVGYLDVSLGAMIGWMRAGEALHGRR 328

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAA 229
                | |: | |  | | |||  | |  :|:  ||||| ::   : ||||
Db        329 TFDATRSPLLNAWMERFAALDAAKAAMPDNNKLVEFVRV---RRAAAA 373


RESULT 8
US-10-437-963-131794
; Sequence 131794, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 131794
;   LENGTH: 306
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_33826C.1.pep
US-10-437-963-131794

  Query Match             45.4%;  Score 542;  DB 16;  Length 306;
  Best Local Similarity   45.1%;  Pred. No. 6.4e-46;
  Matches  102;  Conservative   43;  Mismatches   81;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|:||| |:||:|:| :||| |||| || :|: |  |||||||||||||:||  |  :
Db         12 ELQLLGTWYSPYAMRAKIALGLKGLSYEYIEQDLFGKSELLLKSNPVHKKVPVLIHAGRP 71

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              :||| :::||:|| |   |  :|| : :||| |||| :| |  :    :::     | :
Db         72 VCESRVVLEYVDEAWPGAAPPLLPADPHDRATARFWATYFDSTFFPPWRALMRATTAEQR 131

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |  |   :| :|  | :||:|||:|||| :| ||:  | |: : :| : :     |||
Db        132 AEAFMNAVPQVEVLERAFVECSKGKAFFGGDAVGLVDVVVGGFVVWFKVVDEVAGSSLLD 191

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              | |:|||  ||| | |  ||:  :|:  ||:|  |    | |: ||
Db        192 EAKFPGLAAWAERFLAVDAVREAMPDAGKLLEHYKGFLAKLASPAA 237


RESULT 9
US-10-437-963-117929
; Sequence 117929, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117929
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_2128C.1.pep
US-10-437-963-117929

  Query Match             45.3%;  Score 541;  DB 16;  Length 244;
  Best Local Similarity   49.3%;  Pred. No. 5.9e-46;
  Matches  108;  Conservative   41;  Mismatches   58;  Indels   12;  Gaps    3;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              :|:|| | |||  || :|| |||:::|:::||:  |||||:|||||| ||  |  | :
Db         15 VRVLGGWASPFTNRVVVALKLKGVEHEMLQETVGKKSELLLRSNPVHKKFPVLLHHSKPL 74

Qy         66 CESAIIVEYIDEVW---SNNALSILPQNAYDRANARFWVSYIDDKWLTSLK----SVLAT 118
               || :|||||||||    :  | :|||:: : ||  |||  |:||| |  |:    ||
Db         75 PESLVIVEYIDEVWPASNGGAPAILPRDPHGRAVERFWARYVDDKILPGLRVLRGSVAGD 134

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              :|  | ::        |:::|| | |||:|| |||||:||::||  |||| :|:  | :
Db        135 KDQTAGEM-----STTLQRLEEAFVKCSQGKEYFGGDSIGYLDIALGSFLGWIKAVEKIA 189

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEF 217
              :||:||| | | :||: | | |||  ::|: :|||||
Db        190 GVELLNETKLPILAVWADRFCAHPAVVDVVPDADKLVEF 228


RESULT 10
US-10-437-963-148018
; Sequence 148018, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 148018
;   LENGTH: 235
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_48491C.1.pep
US-10-437-963-148018

  Query Match             45.2%;  Score 540;  DB 16;  Length 235;
  Best Local Similarity   46.8%;  Pred. No. 7.1e-46;
  Matches  108;  Conservative   44;  Mismatches   71;  Indels    8;  Gaps    3;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|:||| | |||||||::||: ||| ||  || |: |||||| |||||||:||  |  |
Db          7 ELKLLGMWTSPFALRVKLALSFKGLSYEYAEEDLSNKSELLLSSNPVHKKVPVLIHNGKP 66

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK----WLTSLKSVLATED 120
              |||| :||:|||| :    : :|| : |:|| |||| :|||||    || :  |:  ||:
Db         67 ICESQVIVQYIDEAFPGAGVPLLPSDPYERAVARFWAAYIDDKLLKSWLQA--SMGKTEE 124
```

```
Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              ::|: |  ::    :  :|  | :||:|| :|||| :|::|:  |: :|::   | :
Db        125 EKAEAL--KETFTAVANLEAAFEECSKGKPFFGGDAVGYLDVTLGALVSWVHAGEALYGM 182

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              :| | |: | |  | : |||  |   :||:   | |:||  | : ||||||
Db        183 RLFDATRSPLLDAWVDRFAALDAAMAVLPDAGTLAEYAKKRQAERAAAAAA 233


RESULT 11
US-10-437-963-152885
; Sequence 152885, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152885
;   LENGTH: 237
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_52892C.1.pep
US-10-437-963-152885

  Query Match             44.8%;  Score 535.5;  DB 16;  Length 237;
  Best Local Similarity   43.9%;  Pred. No. 2e-45;
  Matches  100;  Conservative   46;  Mismatches   81;  Indels    1;  Gaps    1;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|||||  | ||: :||::|| :||| ||  || |: ||: ||:|||||:|:||  || :
Db          9 ELRLLGTWSSPWVIRVRVALGMKGLSYEYTEEDLSSKSDHLLRSNPVHEKVPVLIHGGRP 68

Qy         65 ICESAIIVEYIDEVW-SNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEA 123
              :||| :::||||| | :     :|| : |||| |||| :|::|  :  | | :  :   |
Db         69 VCESLVVLEYIDETWGATGTPQLLPADPYDRATARFWTNYVNDTFFPSWKVLFRSTAAEQ 128

Qy        124 KKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLL 183
              :   |:     :| :|  | :||:|||:||||  | ||:  || | :|:| : :    ||
Db        129 RAEAFKNVVPRVEALERAFGECSKGKAFFGGDDAGLVDVALGSHLVWIKVVDEVAGANLL 188

Qy        184 DETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              || |:|||  ||| | |  ||: ::|:   :::  |    || | | :
Db        189 DEAKFPGLAAWAERFLAVDAVRQVMPDAGDVLKQYKGFLAKWTAGAGS 236


RESULT 12
US-10-437-963-148024
; Sequence 148024, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 148024
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_48497C.1.pep
US-10-437-963-148024

  Query Match             44.8%;  Score 535;  DB 16;  Length 243;
  Best Local Similarity   46.7%;  Pred. No. 2.4e-45;
  Matches  107;  Conservative   45;  Mismatches   73;  Indels    4;  Gaps    2;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|:||| | ||: || : ||: ||| || ||| |: |||||| |||||||:||  |  |
Db         10 ELKLLGLWASPYVLRAKFALSFKGLSYENVEEDLHNKSELLLSSNPVHKKVPVLIHNGKP 69

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLK--SVLATEDDE 122
              |||| |||||:|| : :   |:|| : |||| |||| :||:||:: : :  |:  ||:::
Db         70 ICESQIIVEYVDEAFPDAGESLLPSDPYDRAVARFWAAYINDKFMPAWQKASLGLTEEEK 129

Qy        123 AKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKL 182
              |: :  :|    :| :|  | : |:|| :||||| |::|:  |: : : |  | :  |||
Db        130 AEAV--KQMLAAIENLETAFKELSKGKPFFGGDTAGYLDVTLGAVVGWARAGEVLFGRKL 187

Qy        183 LDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |: | |  | | | |  ||| :||:  :|:|: |:    :|  |||
Db        188 FDATRSPLLAAWMERFVALDAVKAVLPDNAELIEYGKMRMAHYAKLAAA 236


RESULT 13
US-10-437-963-132134
; Sequence 132134, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 132134
;   LENGTH: 236
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_34132C.1.pep
US-10-437-963-132134

  Query Match             44.7%;  Score 534;  DB 16;  Length 236;
  Best Local Similarity   44.9%;  Pred. No. 2.9e-45;
  Matches  105;  Conservative   44;  Mismatches   73;  Indels   12;  Gaps    5;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|:||| | ||:|:||:: |||| | || ||| |  ||:||| |||||| :||  |  :
Db          6 ELKLLGVWSSPYAIRVRVVLNLKSLPYEYVEENLGDKSDLLLASNPVHKSVPVLLHAGRP 65

Qy         65 ICESAIIVEYIDEVWSNNA---LSILPQNAYDRANARFWVSYIDDK----WLTSLKSVLA 117
              : || :||:||||||    |    |::| : |:|| |||| :|:|||    ||  |
Db         66 VNESQVIVQYIDEVWPGGAGGRPSVMPSDPYERAVARFWAAYVDDKVRPAWLAIL---FG 122

Qy        118 TEDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENM 177
              :: :| :     ||   || :|  | :||:|| :|||| :||||:  | :| :    : :
```

```
Db        123 SKTEEERAAAVAQAVAALETLEGAFGECSKGKPFFGGDGVGFVDVVLGGYLGWFTAIDKL 182

Qy        178 NERKLLDETKYPGLTLWAETFAADPAVKGLLP-ETEKLVEFAKILQLKWAAAAA 230
              |:|:|  : | |  | | | |  | ||::| : :||:|| : | |:|:|: |
Db        183 IGRRLIDPARTPALAAWEERFRATDAAKGVVPDDADKLLEFRQTL-LRWSASKA 235


RESULT 14
US-10-437-963-130681
; Sequence 130681, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 130681
;   LENGTH: 336
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(336)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_3281C.1.pep
US-10-437-963-130681

  Query Match             44.6%;  Score 533;  DB 16;  Length 336;
  Best Local Similarity   48.3%;  Pred. No. 5.9e-45;
  Matches  115;  Conservative   37;  Mismatches   70;  Indels   16;  Gaps    4;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              :|::| | |||  || :|| |||:::|:::||:  ||||||:||||||||||  |  | |
Db         98 VRVVGGWASPFMNRVVVALKLKGVEHEMLQETVGKKSELLLRSNPVHKKIPVLLHHGKPI 157

Qy         66 CESAIIVEYIDEVW--SNNALSILPQNAYDRANARFWVSYIDDK------WLTSLKSVLA 117
               || ||| ||||||  |: | :|||:: | ||  |||  |||||       | :| :| :
Db        158 AESLIIVXYIDEVWPASDGAPAILPRDPYCRAVERFWAQYIDDKPDDFALILDNLCAVSS 217

Qy        118 ----TEDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRV 173
                  |  ::    :    |: :|| | |||:|| ||||| ||::||  |||| :|:
Db        218 GDSRVERNRGRRGTVVEMSTALKHLEEAFVKCSQGKKYFGGDKIGYLDIALGSFLGWIKA 277

Qy        174 SENMNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              |     :|||| | | |  ||: | | |||   :|: :|||||    :| ||:  |
Db        278 VEKFAGVELLDEAKVPNLAAWADRFCAHPAVVDAMPDADKLVEFA----VKHAASMKA 331


RESULT 15
US-10-767-701-45399
; Sequence 45399, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof For Plant Improvement
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45399
;   LENGTH: 232
;   TYPE: PRT
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: SORBI-28MAY03-C10423_1.pep
US-10-767-701-45399

  Query Match             44.2%;  Score 528;  DB 16;  Length 232;
  Best Local Similarity   47.7%;  Pred. No. 1.1e-44;
  Matches  106;  Conservative   40;  Mismatches   68;  Indels    8;  Gaps    3;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |:| |:|::| | ||| :|| ||| |||::||:||| :  ||||||:||||||||||  |
Db          1 MSEADVRVIGLWPSPFVIRVLIALKLKGVEYELVEEVVGKKSELLLRSNPVHKKIPVLLH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALS-ILPQNAYDRANARFWVSYIDDKWLTSLKSVLATE 119
                | | || |||:|||||||::  : :|| : | ||  |||  |:|||   ::: :  |:
Db         61 HGKPISESLIIVQYIDEVWSSDTPAFLLPADPYTRAVHRFWAQYVDDKLPPAIRILRGTD 120

Qy        120 ---DDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSEN 176
                  |||     ||    |: :|| | | |:|| |||||::|::||   |:: :::  |
Db        121 GGGKDEAA----EQLSAALQLLEEAFTKLSQGKHYFGGDSVGYLDIALVSYVGWVKAVEK 176

Qy        177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFA 218
              |    |||: | | |  ||:     |||   :|: :| |||:
Db        177 MAGVTLLDKAKVPNLVAWADRLCNHPAVVESIPDADKFVEFS 218


Search completed: October  8, 2004, 09:36:12
Job time : 132 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 8, 2004, 09:15:05 ; Search time 115 Seconds
(without alignments)
636.524 Million cell updates/sec

Title: US-10-088-945A-10
Perfect score: 1195
Sequence: 1 MAERDLRLLGAWFSPFALRV.....KLVEFAKILQLKWAAAAAAK 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1000.5	83.7	234	10	Q9FQF3	Q9fqf3 glycine max
2	638	53.4	227	10	Q9FUS9	Q9fus9 arabidopsis
3	632	52.9	227	10	Q9FUS8	Q9fus8 arabidopsis
4	624	52.2	233	10	Q9SEK1	Q9sek1 papaver som
5	623	52.1	233	10	Q9S7M6	Q9s7m6 papaver som
6	622	52.1	227	10	Q94II0	Q94ii0 arabidopsis
7	617.5	51.7	281	10	Q9M6R4	Q9m6r4 gossypium h
8	596	49.9	234	10	Q9XIF8	Q9xif8 arabidopsis
9	595	49.8	207	10	Q9FUS7	Q9fus7 arabidopsis
10	587	49.1	234	10	Q94K12	Q94k12 arabidopsis
11	573	47.9	233	10	Q9FUE3	Q9fue3 oryza sativ
12	563.5	47.2	226	10	Q9FVI9	Q9fvi9 petroselinu
13	558	46.7	233	10	Q9LQ48	Q9lq48 arabidopsis
14	557.5	46.7	231	10	Q945W4	Q945w4 oryza sativ
15	554	46.4	233	10	Q8LF83	Q8lf83 arabidopsis
16	541	45.3	235	10	Q8S715	Q8s715 oryza sativ
17	541	45.3	244	10	Q945W3	Q945w3 oryza sativ
18	540.5	45.2	230	10	Q8LGS9	Q8lgs9 aegilops ta
19	540	45.2	235	10	Q8S703	Q8s703 oryza sativ
20	539.5	45.1	230	10	O04941	O04941 aegilops ta
21	539	45.1	239	10	Q8RZY9	Q8rzy9 oryza sativ
22	538.5	45.1	255	10	Q8S7S9	Q8s7s9 oryza sativ
23	535.5	44.8	237	10	Q8RZY8	Q8rzy8 oryza sativ
24	535	44.8	243	10	Q945W2	Q945w2 oryza sativ
25	534	44.7	236	10	Q9FUE6	Q9fue6 oryza sativ
26	533	44.6	243	10	Q945X1	Q945x1 oryza sativ
27	531	44.4	239	10	Q8RUJ2	Q8ruj2 oryza sativ
28	530.5	44.4	236	10	Q9FQ99	Q9fq99 zea mays (m
29	528	44.2	234	10	Q8S714	Q8s714 oryza sativ
30	526	44.0	227	10	Q8LBX2	Q8lbx2 arabidopsis
31	525.5	44.0	233	10	Q945W5	Q945w5 oryza sativ
32	523.5	43.8	228	10	Q8LL15	Q8ll15 triticum ae
33	523.5	43.8	229	10	Q9FQ97	Q9fq97 zea mays (m
34	523.5	43.8	233	10	Q8RW01	Q8rw01 triticum ae
35	522.5	43.7	237	10	Q8S712	Q8s712 oryza sativ
36	522	43.7	239	10	Q9FUE4	Q9fue4 oryza sativ
37	521.5	43.6	235	10	Q9SB98	Q9sb98 alopecurus
38	521	43.6	241	10	Q8S718	Q8s718 oryza sativ
39	520.5	43.6	192	10	Q9SY78	Q9sy78 arabidopsis
40	520	43.5	227	10	Q9FUS6	Q9fus6 arabidopsis
41	520	43.5	229	10	Q9FQB7	Q9fqb7 zea mays (m
42	520	43.5	470	10	Q9LFX4	Q9lfx4 arabidopsis
43	519.5	43.5	233	10	Q945W6	Q945w6 oryza sativ
44	519.5	43.5	235	10	Q9ZS20	Q9zs20 alopecurus
45	514.5	43.1	237	10	O65056	O65056 picea maria

ALIGNMENTS

RESULT 1
Q9FQF3
ID Q9FQF3 PRELIMINARY; PRT; 234 AA.
AC Q9FQF3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glutathione S-transferase GST 5 (EC 2.5.1.18).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20532786; PubMed=11080288;
RA McGonigle B., Keeler S.J., Lau S.M.C., Koeppe M.K., O'Keefe D.P.;
RT "A Genomics Approach to the Comprehensive Analysis of the Glutathione
RT S-Transferase Gene Family in Soybean and Maize.";
RL Plant Physiol. 124:1105-1120(2000).
DR EMBL; AF243360; AAG34795.1; -.
DR GO; GO:0004364; F:glutathione transferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
KW Transferase.
SQ SEQUENCE 234 AA; 26463 MW; D29D400227440363 CRC64;

Query Match 83.7%; Score 1000.5; DB 10; Length 234;
Best Local Similarity 84.6%; Pred. No. 6.4e-76;
Matches 192; Conservative 17; Mismatches 17; Indels 1; Gaps 1;

```
Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:|||||||||||:|||||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
```

```
                ||||||||||||| |: | ||||||||||||||| :|||:|| |||:||| ||||||
Db          66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy         125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
               | ||||||| ||::|||||| ||||||||||:|||:|||||||||::|| | |: |||||
Db         125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy         185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
               | |:|||| |||||||||||||:||||:|||||||||||||| |||||
Db         185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 2
Q9FUS9
ID   Q9FUS9      PRELIMINARY;      PRT;   227 AA.
AC   Q9FUS9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase.
GN   GST29 OR AT1G10360.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wagner U., Mauch F.;
RT   "Analysis of the glutathione S-transferase family in Arabidopsis
RT   thaliana.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA   Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA   Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
RA   Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
RA   Ecker J.R., Theologis A.;
RT   "Arabidopsis Open Reading Frame (ORF) Clones.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288190; AAG30139.1; -.
DR   EMBL; BT000940; AAN41340.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   227 AA;  25911 MW;  2996DA12B9BF590B CRC64;

  Query Match             53.4%;  Score 638;  DB 10;  Length 227;
  Best Local Similarity   54.3%;  Pred. No. 1.6e-45;
  Matches  120;  Conservative   34;  Mismatches   67;  Indels    0;  Gaps    0;

Qy           1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
               ||  |::|:|:| | : :| :|||:|| : || ::||   |||||||||||||:||  |
Db           1 MATEDVKLIGSWASVYVMRARIALHLKSISYEFLQETYGSKSELLLKSNPVHKKMPVLIH 60

Qy          61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               || :||| ||| |||| |::  |||| : |||| |||| :||||:|  |::|:|  :
Db          61 ADKPVCESNIIVHYIDEAWNSSGPSILPSHPYDRAIARFWAAYIDDQWFISVRSILTAQG 120

Qy         121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
               || ||    | ||  : :|: || ||:|| :| || ||::||  |||| : || |
Db         121 DEEKKAAIAQVEERTKLLEKAFNDCSQGKPFFNGDHIGYLDIALGSFLGWWRVVELDANH 180

Qy         181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKIL 221
               | ||||| | |  ||| |  ||||| ::||  || |||: |
Db         181 KFLDETKTPSLVKWAERFCDDPAVKPIMPEITKLAEFARKL 221


RESULT 3
```

```
Q9FUS8
ID   Q9FUS8      PRELIMINARY;      PRT;   227 AA.
AC   Q9FUS8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase.
GN   GST30.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wagner U., Mauch F.;
RT   "Analysis of the glutathione S-transferase family in Arabidopsis
RT   thaliana.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288191; AAG30140.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   227 AA;  25307 MW;  6A1743B849937442 CRC64;

  Query Match             52.9%;  Score 632;  DB 10;  Length 227;
  Best Local Similarity   55.3%;  Pred. No. 5e-45;
  Matches  121;  Conservative   32;  Mismatches   66;  Indels    0;  Gaps    0;

Qy           1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
               ||  |::|:||| ||| :| :|||||| : || ::||   |||||||||||||||  |
Db           1 MASSDVKLIGAWASPFVMRPRIALNLKSVPYEFLQETFGSKSELLLKSNPVHKKIPVLLH 60

Qy          61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               || : || |||||||: ||::  |||| : |||| |||| :|||:||  :|:  |
Db          61 ADKPVSESNIIVEYIDDTWSSSGPSILPSDPYDRAMARFWAAYIDEKWFVALRGFLKAGG 120

Qy         121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
               :| ||    | ||    :|: |  ||:|| :| || ||::||  | ||:::||:|
Db         121 EEEKKAVIAQLEEGNAFLEKAFIDCSKGKPFFNGDNIGYLDIALGCFLAWLRVTELAVSY 180

Qy         181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
               |:||| | | |: ||| |  ||||| ::||| || ||||
Db         181 KILDEAKTPSLSKWAENFCNDPAVKPVMPETAKLAEFAK 219


RESULT 4
Q9SEK1
ID   Q9SEK1      PRELIMINARY;      PRT;   233 AA.
AC   Q9SEK1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase 3 (EC 2.5.1.18).
OS   Papaver somniferum (Opium poppy).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC   Papaveraceae; Papaver.
OX   NCBI_TaxID=3469;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Marianne;
RA   Facchini P.J., Yu M.;
RT   "Molecular cloning and characterization of a glutathione S-transferase
RT   gene family from opium poppy.";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF118926; AAF22519.1; -.
DR   GO; GO:0004364; F:glutathione transferase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
```

```
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   233 AA;  26017 MW;  C5A1376BC0D83B40 CRC64;

  Query Match             52.2%;  Score 624;  DB 10;  Length 233;
  Best Local Similarity   54.0%;  Pred. No. 2.4e-44;
  Matches  116;  Conservative   39;  Mismatches   60;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::|||||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKIPVMIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|| : ||::|| || :: :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGEEKIGYIDIAFGCYIGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              ||| |||| ||| | ||  || ::|||: |:|||
Db        188 ETKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 5
Q9S7M6
ID   Q9S7M6      PRELIMINARY;      PRT;   233 AA.
AC   Q9S7M6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase 2 (EC 2.5.1.18).
OS   Papaver somniferum (Opium poppy).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC   Papaveraceae; Papaver.
OX   NCBI_TaxID=3469;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. MARIANNE;
RA   Facchini P.J., Yu M.;
RT   "Molecular cloning and characterization of a glutathione S-transferase
RT   gene family from opium poppy.";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF118925; AAF22518.1; -.
DR   EMBL; AF118924; AAF22517.1; -.
DR   GO; GO:0004364; F:glutathione transferase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   233 AA;  25945 MW;  72FF3137BC295CDD CRC64;

  Query Match             52.1%;  Score 623;  DB 10;  Length 233;
  Best Local Similarity   54.0%;  Pred. No. 2.9e-44;
  Matches  116;  Conservative   39;  Mismatches   60;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::||:||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKMPVLIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|||: ||:||| || :: :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGGEKIGYVDIAFGCYVGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              | | |||| ||| | ||  || ::|||: |:|||
Db        188 EEKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 6
Q94II0
ID   Q94II0      PRELIMINARY;      PRT;   227 AA.
AC   Q94II0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase.
GN   ERD9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RA   Shinozaki K., Yamaguchi-shinozaki K., Takahashi S.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=94355652; PubMed=8075396;
RA   Kiyosue T., Yamaguchi-shinozaki K., Shinozaki K.;
RT   "Cloning of cDNAs for genes that are early-responsive to dehydration
RT   stress (ERDs) in Arabidopsis thaliana L.: identification of three ERDs
RT   as HSP cognate genes.";
RL   Plant Mol. Biol. 25:791-798(1994).
DR   EMBL; AB039930; BAB63917.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   227 AA;  25246 MW;  70A6FD56B98C9BB2 CRC64;

  Query Match             52.1%;  Score 622;  DB 10;  Length 227;
  Best Local Similarity   54.8%;  Pred. No. 3.4e-44;
  Matches  120;  Conservative   32;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||  |::|:||| ||| :| :|||||| : || ::||   ||||||||||||||||  |
Db          1 MASSDVKLIGAWASPFVMRPRIALNLKSVPYEFLQETFGSKSELLLKSNPVHKKIPVLLH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
              || : || |||||||: ||::  |||| : |||| |||| :|||:||  :|:  |
Db         61 ADKPVSESNIIVEYIDDTWSSSGPSILPSDPYDRAMARFWAAYIDEKWFVALRGFLKAGG 120

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              :| ||    | ||    :|: |  ||:||  | || ||:::|  | ||:::||:|
Db        121 EEEKKAVIAQLEEGNAFLEKAFIDCSKGKPSFNGDNIGYLNIALGCFLAWLRVTELAVSY 180

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              |:||| | | |: ||| |  ||||| ::||| || ||||
Db        181 KILDEAKTPSLSKWAENFCNDPAVKPVMPETAKLAEFAK 219


RESULT 7
Q9M6R4
ID   Q9M6R4      PRELIMINARY;      PRT;   281 AA.
AC   Q9M6R4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase.
OS   Gossypium hirsutum (Upland cotton).
```

```
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=3635;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fiber;
RA   Kang W.H., Yamamoto E., Allen D.R.;
RT   "Cloning and characterization of glutathione S-transferase (GST) gene
RT   from cotton plants (Gossypium hirsutum L. cv. Coker 312).";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF159229; AAF29773.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   281 AA;  31803 MW;  75B09DD3C866BFDC CRC64;

  Query Match             51.7%;  Score 617.5;  DB 10;  Length 281;
  Best Local Similarity   53.2%;  Pred. No. 1.1e-43;
  Matches  118;  Conservative   42;  Mismatches   59;  Indels    3;  Gaps    2;

Qy          2 AERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVFFH 60
              :: :::::|| | ||| :||:||||:| : || ::| |   ||||||||||||:|:||  |
Db         53 SDAEVKVLGTWASPFVMRVRIALNIKSVAYEFLQERLWEGKSELLLKSNPVHRKVPVLIH 112

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
              ||  |||| |||:|||||   :: ||| : ::|| |||| :|:|||| ||:::   |
Db        113 GDDTICESLIIVQYIDEVWP--SVPILPSDPHERAAARFWAAYLDDKWFPSLRAIGMAEG 170

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              ::|:|    | || |  :|| | |||:|:| || | ||::|| || || ::||:| |:
Db        171 EDARKAAIGQVEEGLMLLEEAFGKCSQGQALFGKDQIGYLDITFGCFLGWLRVTEKMSGI 230

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQ 222
              |||:|   | |  ||  |  | ||| ::|||||| ||||:|:
Db        231 KLLNEINTPALLKWANRFCNDAAVKDVMPETEKLAEFAKMLR 272


RESULT 8
Q9XIF8
ID   Q9XIF8      PRELIMINARY;      PRT;   234 AA.
AC   Q9XIF8;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   F23H11.1 protein (T30E16.30).
GN   F23H11.1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA   Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA   Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA   Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA   Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Chao Q., Brooks S., Buehler E., Dunn P., Khan S., Kim C., Shinn P.,
RA   Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A., Hansen N.F.,
RA   Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S., Luros S.,
RA   Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G., Davis R.W.,
RA   Federspiel N.A., Theologis A., Ecker J.R.;
RT   "Genomic sequence for Arabidopsis thaliana BAC T30E16 from chromosome
RT   I.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Ecker J.R.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Ecker J.R.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA   Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA   Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA   Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA   Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA   Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA   Theologis A., Ecker J.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007258; AAD39312.1; -.
DR   EMBL; AC009317; AAF79760.1; -.
DR   PIR; F96620; F96620.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
SQ   SEQUENCE   234 AA;  26506 MW;  F744F7C93A81E843 CRC64;

  Query Match             49.9%;  Score 596;  DB 10;  Length 234;
  Best Local Similarity   49.6%;  Pred. No. 5.4e-42;
  Matches  115;  Conservative   46;  Mismatches   69;  Indels    2;  Gaps    2;

Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | |: :::||| |:||:|:| :||| || :||: ||| |   ||||||||||||:||
Db          1 MGEKEEVKLLGVWYSPYAIRPKIALRLKSVDYDYVEENLFGSKSELLLKSNPVHKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
              | :| | ||  ||||||| |:::| |||| : |||| ||||  ::|:||  :|:    |
Db         61 LHNNKPIVESLNIVEYIDETWNSSAPSILPSHPYDRALARFWSDFVDNKWFPALRMAAIT 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              : ::||    |: || | ::|: |   |:|| :|||: |||:|| ||||:  ::  |
Db        121 KSEDAKAKAMEEVEEGLLQLEDAFVSISKGKPFFGGEAIGFMDICFGSFVVLLKAREKFK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              |||||:| | |  ||: | :|  || : || ||: || : |:::  :||:
Db        181 AEKLLDESKTPSLCKWADRFLSDETVKNVAPEIEKVAEFLQELEVRAQSAAS 232


RESULT 9
Q9FUS7
ID   Q9FUS7      PRELIMINARY;      PRT;   207 AA.
AC   Q9FUS7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase.
GN   GST30B.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wagner U., Mauch F.;
RT   "Analysis of the glutathione S-transferase family in Arabidopsis
RT   thaliana.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288192; AAG30141.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
```

```
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   207 AA;  23049 MW;  98149A1E4FAABEB8 CRC64;

  Query Match           49.8%;  Score 595;  DB 10;  Length 207;
  Best Local Similarity   52.5%;  Pred. No. 5.6e-42;
  Matches  115;  Conservative   29;  Mismatches   55;  Indels   20;  Gaps    1;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||  |::|:||| ||| :| :|||||| : || ::||   |||||||||||||||||  |
Db          1 MASSDVKLIGAWASPFVMRPRIALNLKSVPYEFLQETFGSKSELLLKSNPVHKKIPVLLH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               || : || ||||||| ||::  |||| : |||| |||| :|||:||  :|:  |
Db         61 ADKPVSESNIIVEYIDDTWSSSGPSILPSDPYDRAMARFWAAYIDEKWFVALRGFL---- 116

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
                          |    |  ||:|| :| || ||::||  | ||:::||:| .
Db        117 ----------------KAGGAFIDCSKGKPFFNGDNIGYLDIALGCFLAWLRVTELAVSY 160

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              |:||| | | |: ||| |  ||||| ::||| || ||||
Db        161 KILDEAKTPSLSKWAENFCNDPAVKPVMPETAKLAEFAK 199


RESULT 10
Q94K12
ID   Q94K12      PRELIMINARY;      PRT;   234 AA.
AC   Q94K12;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Similar to glutathione S-transferase.
GN   T30E16.30.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA   Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA   Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA   Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA   Ecker J., Theologis A., Davis R.W.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF370480; AAK43857.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   234 AA;  26490 MW;  E8A4F7D6DA81E843 CRC64;

  Query Match           49.1%;  Score 587;  DB 10;  Length 234;
  Best Local Similarity   49.1%;  Pred. No. 3.1e-41;
  Matches  114;  Conservative   46;  Mismatches   70;  Indels    2;  Gaps    2;

Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | |: :::|||  |:||:|:| :||| || :||: ||| |   ||||||||||||||:||
Db          1 MGEKEEVKLLGVWYSPYAIRPKIALRLKSVDYDYVEENLFGSKSELLLKSNPVHKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
                | :| | ||  ||||||| |:::| |||| : |||| ||||  ::|:||  :|:    |
Db         61 LHNNKPIVESLNIVEYIDETWNSSAPSILPSHPYDRALARFWSDFVDNKWFPALRMAAIT 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
               : ::||    |: || | ::|: |   |:|| :|||: |||:|| ||||:  ::  |
Db        121 KSEDAKAKAMEEVEEGLLQLEDAFVSISKGKPFFGGEAIGFMDICFGSFVVLLKAREKFK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
               ||| |:| | |  ||: | :|  || : || ||: || : |:::  :||:
Db        181 AEKLLVESKTPSLCKWADRFLSDETVKNVAPEIEKVAEFLQELEVRAQSAAS 232


RESULT 11
Q9FUE3
ID   Q9FUE3      PRELIMINARY;      PRT;   233 AA.
AC   Q9FUE3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative glutathione S-transferase.
GN   OSJNBA0034L04.7 OR OSJNBB0038A07.10.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Soranzo N., Frova C., Rizzardi R., Sari-Gorla M.;
RT   "Characterization of EST clones encoding for glutathione S-
RT   transferases in Oryza sativa (L.).";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA   Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA   Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,
RA   VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA   Salzberg S.L., White O., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBa0034L04 genomic sequence.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA   Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
RA   Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
RA   Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
RA   Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBb0038A07 genomic sequence.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
RT   chromosome 10.";
RL   Science 300:1566-1569(2003).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF309379; AAG32472.1; -.
DR   EMBL; AC091680; AAM12325.1; -.
DR   EMBL; AC113948; AAM94544.1; -.
DR   EMBL; AE017114; AAP54745.1; -.
DR   Gramene; Q9FUE3; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   233 AA;  25332 MW;  C80416E1EA81CBD5 CRC64;

  Query Match           47.9%;  Score 573;  DB 10;  Length 233;
```

```
  Best Local Similarity   47.6%;  Pred. No. 4.6e-40;
  Matches  107;  Conservative   43;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|:||| | |||||| ::||: ||| || ||| |  |||||| |||||||:||  |  |
Db          7 ELKLLGMWASPFALRAKLALSFKGLSYEYVEEDLKNKSELLLTSNPVHKKVPVLIHNGKP 66

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              :||| :||:|||| : :  : :|| : |||| |||| :|||| | |       : :: |
Db         67 VCESQVIVQYIDEAFPDAGVPLLPSDPYDRAVARFWAAYIDDKLLKSWLQASMCKTEQEK 126

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
                ::    :  :|  | :||:|| :|||| :|:||:  |: : |:|| | ::  :| |
Db        127 AAAMKETFAAVANLEAAFKECSKGKPFFGGDAVGYVDVTLGAVIGFVRVGEAVHGMRLFD 186

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAA 229
              :: | |  | : |||  | | :||:| :| |:||: | :|||||
Db        187 ASRSPLLDAWLDRFAALDAAKAVLPDTGRLAEYAKMKQAEWAAAA 231


RESULT 12
Q9FVI9
ID   Q9FVI9      PRELIMINARY;      PRT;   226 AA.
AC   Q9FVI9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Petroselinum crispum (Parsley) (Petroselinum hortense).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   campanulids; Apiales; Apiaceae; Apioideae; apioid superclade;
OC   Apium clade; Petroselinum.
OX   NCBI_TaxID=4043;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20496825; PubMed=11041888;
RA   Loyall L., Uchida K., Braun S., Furuya M., Frohnmeyer H.;
RT   "Glutathione and a UV light-induced glutathione S-transferase are
RT   involved in signaling to chalcone synthase in cell cultures.";
RL   Plant Cell 12:1939-1950(2000).
DR   EMBL; AF177944; AAG09294.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   226 AA;  25165 MW;  D1453F3C5C845B5B CRC64;

  Query Match             47.2%;  Score 563.5;  DB 10;  Length 226;
  Best Local Similarity   52.4%;  Pred. No. 2.8e-39;
  Matches  111;  Conservative   33;  Mismatches   67;  Indels    1;  Gaps    1;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              ::||| | |||  |||||||:| : ||::|| |  ||:|||:|||||||:||  |||| |
Db          8 VKLLGFWASPFVNRVQIALNIKSISYELIEENLRAKSQLLLESNPVHKKVPVLIHGDKPI 67

Qy         66 CESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAKK 125
                || :||||||| |::   |||| : : || |||| :|| |||   |       | :|||
Db         68 SESLVIVEYIDEAWTHGP-SILPSDPFARATARFWAAYITDKWWPLLTEYRTAEGEEAKA 126

Qy        126 LHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLDE 185
                |:  | |  :|| | | |:|| ||||| ||::|:  || |:::|| |        ||:
Db        127 AAKEKFLEGLVLLEEAFVKSSQGKGYFGGDNIGYLDVVLGSLLAWLRVREVTQGVAFLDK 186

Qy        186 TKYPGLTLWAETFAADPAVKGLLPETEKLVEF 217
              | | |  ||: | ::|||  :||:|||:|:|
Db        187 AKAPELAAWADRFTSNPAVVDVLPKTEKVVDF 218


RESULT 13
Q9LQ48
ID   Q9LQ48      PRELIMINARY;      PRT;   233 AA.
AC   Q9LQ48;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   T30E16.25.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chao Q., Brooks S., Buehler E., Dunn P., Khan S., Kim C., Shinn P.,
RA   Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A., Hansen N.F.,
RA   Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S., Luros S.,
RA   Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G., Davis R.W.,
RA   Federspiel N.A., Theologis A., Ecker J.R.;
RT   "Genomic sequence for Arabidopsis thaliana BAC T30E16 from chromosome
RT   I.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ecker J.R.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Ecker J.R.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA   Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA   Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA   Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA   Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA   Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA   Theologis A., Ecker J.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC009317; AAF79758.1; -.
DR   PIR; D96620; D96620.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
SQ   SEQUENCE   233 AA;  26535 MW;  8722CA2CE7AEC09D CRC64;

  Query Match             46.7%;  Score 558;  DB 10;  Length 233;
  Best Local Similarity   48.4%;  Pred. No. 8.3e-39;
  Matches  106;  Conservative   42;  Mismatches   69;  Indels    2;  Gaps    2;

Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | || :::||| |:|| :| :||| || :||: ||| |   ||||||||||: ||:||
Db          1 MGEREEVKLLGTWYSPVVIRAKIALRLKSVDYDYVEEDLFGSKSELLLKSNPIFKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
                | | :| |  ||||||| |:::  |||| : |||| ||||  ::||||| :| : :
Db         61 IHNTKPVCVSLNIVEYIDETWNSSGSSILPSHPYDRALARFWSVFVDDKWLPTLMAAVVA 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              : :|||    |: || | ::|  |   |:||::|||:||||:||  ||||  ::  | :
Db        121 KSEEAKAKGMEEVEEGLLQLEAAFIALSKGKSFFGGETIGFIDICLGSFLVLLKAREKLK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEF 217
              |:||| | | |  ||  | ::  || ::|: :|: :|
Db        181 NEKILDELKTPSLYRWANQFLSNEMVKNVVPDIDKVAKF 219


RESULT 14
Q945W4
ID   Q945W4      PRELIMINARY;      PRT;   231 AA.
AC   Q945W4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative glutathione S-transferase.
GN   OSJNBA0034L04.6 OR OSJNBB0038A07.9.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare; TISSUE=Shoot;
RA   Soranzo N., Sari-Gorla M.;
RT   "Characterization of EST clones encoding glutathione S-transferases in
RT   Oryza sativa (L.).";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA   Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA   Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,
RA   VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA   Salzberg S.L., White O., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBa0034L04 genomic sequence.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA   Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
RA   Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
RA   Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
RA   Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBb0038A07 genomic sequence.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
RT   chromosome 10.";
RL   Science 300:1566-1569(2003).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF402801; AAK98542.1; -.
DR   EMBL; AC091680; AAM12308.1; -.
DR   EMBL; AC113948; AAM94539.1; -.
DR   EMBL; AE017114; AAP54749.1; -.
DR   Gramene; Q945W4; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   231 AA;  25672 MW;  86CA4146CFE57A36 CRC64;

  Query Match             46.7%;  Score 557.5;  DB 10;  Length 231;
  Best Local Similarity   47.8%;  Pred. No. 9e-39;
  Matches  109;  Conservative   46;  Mismatches   62;  Indels   11;  Gaps    4;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              |:||| | ||:  ||::||: ||| || ||| |: |||||| :||||||:||  |  | |
Db          9 LKLLGLWVSPYTHRVKLALSFKGLSYEYVEEDLSNKSELLLSTNPVHKKVPVLIHNGKPI 68

Qy         66 CESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK----WLTSLKSVLATEDD 121
              ||| :||:|:|| : |: :|:|| ::|||| |||| :||:||    || |  |:  ||::
Db         69 CESQVIVQYLDEEFPNSGVSLLPSDSYDRAIARFWAAYINDKLMPSWLQS--SMGKTEEE 126

Qy        122 EAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERK 181
              |: |  :|  | :  :|  | :||:|| :|||||:|::|:  |: : ::|  | :: |:
Db        127 RAEAL--KQTLEAVANLETAFKECSKGKPFFGGDTVGYLDVSLGAMIGWMRAGEALHGRR 184

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAA 229
                | |: | |  | | |||  | |  :|:  ||||| ::   : ||||
Db        185 TFDATRSPLLNAWMERFAALDAAKAAMPDNNKLVEFVRV---RRAAAA 229


RESULT 15
Q8LF83
ID   Q8LF83      PRELIMINARY;      PRT;   233 AA.
AC   Q8LF83;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Glutathione S-transferase, putative.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 0:0-0(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY084992; AAM61551.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   233 AA;  26565 MW;  A6F425E6EF10B774 CRC64;

  Query Match             46.4%;  Score 554;  DB 10;  Length 233;
  Best Local Similarity   47.9%;  Pred. No. 1.8e-38;
  Matches  105;  Conservative   43;  Mismatches   69;  Indels    2;  Gaps    2;

Qy          1 MAERD-LRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | ||: ::||| |:||  :| :| | || :||: ||| |   ||||||||||:||:||
Db          1 MGEREQVKLLGTWYSPVVIRAKIDLRLKSVDYDYVEENLFGSKSELLLKSNPVYKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
                | | :| |  |||||||| |:::  |||| : |||| ||||  ::||||| :| : :
Db         61 IHNTKPVCVSLNIVEYIDETWNSSGSSILPSHPYDRALARFWSVFVDDKWLPTLMAAVVA 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              : :|||    |: || | :::  |   |:||::|||:||||:||  ||||  ::  | :
Db        121 KSEEAKAKGMEEVEEGLLQLDAAFIALSKGKSFFGGETIGFIDICLGSFLVLLKAREKLK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEF 217
              |:||| | | |  ||  | ::  || ::|: :|: :|
Db        181 NEKILDELKTPSLYRWANQFLSNEMVKNVVPDIDKVAKF 219


Search completed: October  8, 2004, 09:24:14
Job time : 119 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 8, 2004, 09:05:05 ; Search time 25 Seconds
(without alignments)
483.211 Million cell updates/sec

Title: US-10-088-945A-10
Perfect score: 1195
Sequence: 1 MAERDLRLLGAWFSPFALRV.....KLVEFAKILQLKWAAAAAAK 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	429.5	35.9	225	1	GTX6_SOYBN	P32110 glycine max
2	419.5	35.1	236	1	GTX2_MAIZE	P50472 zea mays (m
3	410.5	34.4	221	1	GTXC_TOBAC	P49332 nicotiana t
4	407.5	34.1	221	1	GTX4_TOBAC	Q03666 nicotiana t
5	402	33.6	220	1	GTXA_TOBAC	P25317 nicotiana t
6	399	33.4	224	1	GTXA_ARATH	P46421 arabidopsis
7	392	32.8	223	1	GTX1_TOBAC	Q03662 nicotiana t
8	378	31.6	219	1	LGUL_SOYBN	P46417 glycine max
9	373.5	31.3	223	1	GTX2_TOBAC	Q03663 nicotiana t
10	372	31.1	217	1	GTX1_SOLTU	P32111 solanum tub
11	371.5	31.1	223	1	GTX3_TOBAC	Q03664 nicotiana t
12	361.5	30.3	219	1	GTX1_NICPL	P50471 nicotiana p
13	203.5	17.0	254	1	GTXC_ORYSA	Q06398 oryza sativ
14	193	16.2	241	1	GTO1_PIG	Q9n1f5 sus scrofa
15	192	16.1	241	1	GTO1_HUMAN	P78417 homo sapien
16	180.5	15.1	257	1	LIGF_PSEPA	P30347 pseudomonas
17	170	14.2	241	1	GTO1_RAT	Q9z339 rattus norv
18	166.5	13.9	240	1	GTO1_MOUSE	O09131 mus musculu
19	166	13.9	213	1	GTH2_TOBAC	P46440 nicotiana t
20	164	13.7	213	1	GSTZ_WHEAT	O04437 triticum ae
21	164	13.7	221	1	GTZ1_DIACA	P28342 dianthus ca
22	163	13.6	213	1	GTH1_TOBAC	P30109 nicotiana t
23	155.5	13.0	212	1	GTH_HYOMU	P46423 hyoscyamus
24	154	12.9	221	1	GTZ1_ARATH	Q9zvq3 arabidopsis
25	152.5	12.8	243	1	GTO2_HUMAN	Q9h4y5 homo sapien
26	148.5	12.4	214	1	MAAI_CAEEL	Q18938 caenorhabdi
27	146	12.2	223	1	GTZ2_ARATH	Q9zvq4 arabidopsis
28	140.5	11.8	215	1	MAAI_VIBCH	Q9ksb2 vibrio chol
29	140	11.7	211	1	SSPA_ECOLI	P05838 escherichia
30	139	11.6	250	1	YK67_CAEEL	P34345 caenorhabdi
31	134.5	11.3	221	1	GTH3_MAIZE	P04907 zea mays (m
32	134	11.2	208	1	GTH1_ARATH	P42760 arabidopsis
33	132.5	11.1	253	1	YKJ3_CAEEL	P34277 caenorhabdi
34	131	11.0	225	1	GSTZ_EUPES	P57108 euphorbia e
35	130	10.9	212	1	GTHC_ARATH	Q9slm6 arabidopsis
36	128.5	10.8	209	1	GTHB_ARATH	Q9sry5 arabidopsis
37	127	10.6	145	1	GTZ2_DIACA	Q03425 dianthus ca
38	125.5	10.5	225	1	GTA_PLEPL	P30568 pleuronecte
39	124	10.4	212	1	SSPA_HAEIN	P45207 haemophilus
40	123	10.3	211	1	GTH4_ARATH	P46422 arabidopsis
41	121.5	10.2	215	1	GTH6_ARATH	Q96266 arabidopsis
42	120	10.0	213	1	GTH1_MAIZE	P12653 zea mays (m
43	118	9.9	246	1	MAA1_DROME	Q9vhd3 drosophila
44	115	9.6	209	1	GT_HAEIN	P44521 haemophilus
45	115	9.6	216	1	MAAI_MOUSE	Q9wvl0 mus musculu

ALIGNMENTS

RESULT 1
GTX6_SOYBN
ID GTX6_SOYBN STANDARD; PRT; 225 AA.
AC P32110;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Probable glutathione S-transferase (EC 2.5.1.18) (Heat shock protein
DE 26A) (G2-4).
GN HSP26-A.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88216585; PubMed=2835661;
RA Czarnecka E., Nagao R.T., Key J.L., Gurley W.B.;
RT "Characterization of Gmhsp26-A, a stress gene encoding a divergent
RT heat shock protein of soybean: heavy-metal-induced inhibition of
RT intron processing.";
RL Mol. Cell. Biol. 8:1113-1122(1988).
RN [2]
RP SEQUENCE OF 86-225 FROM N.A.
RX MEDLINE=88198196; PubMed=3360788;
RA Hagen G., Uhrhammer N., Guilfoyle T.J.;
RT "Regulation of expression of an auxin-induced soybean sequence by
RT cadmium.";
RL J. Biol. Chem. 263:6442-6446(1988).
CC -!- FUNCTION: May play a role in the cellular response to stress.
CC -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC -!- INDUCTION: By heat shock and auxin; by heavy metals like cadmium,
CC silver and copper.
CC -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC ----------

CC ----------
DR EMBL; M20363; AAA33973.1; -.
DR EMBL; J03197; AAA33943.1; -.
DR PIR; A33654; A33654.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
KW Transferase; Heat shock.
FT CONFLICT 216 216 A -> E (IN REF. 2).
SQ SEQUENCE 225 AA; 26011 MW; 9E89AC46A312E240 CRC64;

Query Match 35.9%; Score 429.5; DB 1; Length 225;

```
  Best Local Similarity   40.7%;  Pred. No. 1.5e-27;
  Matches   94;  Conservative   46;  Mismatches   80;  Indels   11;  Gaps    5;

Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGD 62
              : |:|||   |||  |||||| |||::|: :|| |  ||:|||| ||||||:||| | :
Db          5 QEDVKLLGIVGSPFVCRVQIALKLKGVEYKFLEENLGNKSDLLLKYNPVHKKVPVFVHNE 64

Qy         63 KVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDE 122
              : | || :|||||| | ||   ||| : | || ||||  :|||| : ::   : | |::
Db         65 QPIAESLVIVEYIDETWKNN--PILPSDPYQRALARFWSKFIDDKIVGAVSKSVFTVDEK 122

Qy        123 AKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSF-IRVSENMNERK 181
              :: : |:  | |: :|     : | :|||:  | |||    |::| | : : :   :
Db        123 EREKNVEETYEALQFLENEL----KDKKFFGGEEFGLVDIA-AVFIAFWIPIFQEIAGLQ 177

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAK 232
              |    |:| |  |:: |   | |  :||  : |  : |    :: : :|:|
Db        178 LFTSEKFPILYKWSQEFLNHPFVHEVLPPRDPLFAYFK---ARYESLSASK 225


RESULT 2
GTX2_MAIZE
ID   GTX2_MAIZE     STANDARD;      PRT;   236 AA.
AC   P50472;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable glutathione S-transferase BZ2 (EC 2.5.1.18) (Bronze-2
DE   protein).
GN   BZ2.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Seedling;
RX   MEDLINE=93005645; PubMed=1967051;
RA   Nash J., Luehrsen K.R., Walbot V.;
RT   "Bronze-2 gene of maize: reconstruction of a wild-type allele and
RT   analysis of transcription and splicing.";
RL   Plant Cell 2:1039-1049(1990).
RN   [2]
RP   ERRATUM.
RA   Nash J., Luehrsen K.R., Walbot V.;
RL   Plant Cell 3:103-103(1991).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- PATHWAY: Anthocyanin biosynthesis.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U14599; AAA50245.1; -.
DR   MaizeDB; 64140; -.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   236 AA;  25563 MW;  BD7C087F60FA6E9A CRC64;

  Query Match           35.1%;  Score 419.5;  DB 1;  Length 236;
  Best Local Similarity   40.4%;  Pred. No. 1e-26;
  Matches   92;  Conservative   48;  Mismatches   79;  Indels    9;  Gaps    5;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNP-KSELLLKSNPVHKKIPVFFHGD-K 63
              :|:||   |||  | ::||:|:|: ||:::| | | ||: || :|||: ||||    | :
Db          1 MRVLGGEVSPFTARARLALDLRGVAYELLDEPLGPKKSDRLLAANPVYGKIPVLLLPDGR 60

Qy         64 VICESAIIVEYIDEVWSNNALS-----ILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
              |||||:||:||::|    :  :     :|| : |:||  ||| ::||||:  :| :|
Db         61 AICESAVIVQYIEDVARESGGAEAGSLLLPDDPYERAMHRFWTAFIDDKFWPALDAVSLA 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYF-GGDTI-GFVDIGFGSFLSFIRVSEN 176
                 |:    |    |  :|| |   | |:|:|  |||    | :|:  | ||  :|  |
Db        121 PTPGARAQAAEDTRAALSLLEEAFKDRSNGRAFFSGGDAAPGLLDLALGCFLPALRACER 180

Qy        177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLK 224
              ::   |:| :  | |  |:: ||| || | :||:|||:|:| : ||::
Db        181 LHGLSLIDASATPLLDGWSQRFAAHPAAKRVLPDTEKVVQFTRFLQVQ 228


RESULT 3
GTXC_TOBAC
ID   GTXC_TOBAC     STANDARD;      PRT;   221 AA.
AC   P49332;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Probable glutathione S-transferase parC (EC 2.5.1.18) (Auxin-regulated
DE   protein parC).
GN   PARC.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Xanthi NC; TISSUE=Leaf mesophyll;
RA   Takahashi Y., Nagata T.;
RT   "Differential expression of an auxin-regulated gene, parC, and a
RT   novel related gene, C-7 from tobacco mesophyll protoplasts in
RT   response to external stimuli and plant tissues.";
RL   Plant Cell Physiol. 33:779-787(1992).
CC   -!- FUNCTION: Conjugation of reduced glutathione to a wide number of
CC       exogenous and endogenous hydrophobic electrophiles.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- TISSUE SPECIFICITY: Abundant in seedlings and roots. It is also
CC       found in the shoot tips, flowers and leaves.
CC   -!- INDUCTION: By auxin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. Phi family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X64398; CAA45740.1; -.
DR   PIR; S19185; S19185.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase; Multigene family.
SQ   SEQUENCE   221 AA;  25743 MW;  18D27C69857EB3AE CRC64;

  Query Match           34.4%;  Score 410.5;  DB 1;  Length 221;
  Best Local Similarity   41.4%;  Pred. No. 5e-26;
  Matches   94;  Conservative   31;  Mismatches   91;  Indels   11;  Gaps    3;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||  :: ||  | | | :|::|||  | : ||  :| |  || ||| :||:|||||  |
Db          1 MANEEVILLDFWPSMFGMRLRIALAEKEIKYEYKQEDLRNKSPLLLQMNPIHKKIPVLIH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT-- 118
                | |||| | ||||:||| : | |:|| : |||| ||||  ||| |     : : ||
Db         61 NGKPICESIIAVEYIEEVWKDKAPSLLPSDPYDRAQARFWADYIDKKLYDFGRKLWATKG 120
```

```
Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              |: || |  | :  :|||          : ||||:: |||||    | |:    |
Db        121 EEQEAAKKDFIECLKVLE-------GALGDRPYFGGESFGFVDIALIGFYSWFYAYETFG 173

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKW 225
                  | : |    ||:      :|   ||:  |::|| |:|: |:
Db        174 --NFSTEAECPKFVAWAKRCMQRESVAKSLPDQPKVLEFVKVLRQKF 218


RESULT 4
GTX4_TOBAC
ID   GTX4_TOBAC     STANDARD;      PRT;   221 AA.
AC   Q03666;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Probable glutathione S-transferase (EC 2.5.1.18) (Auxin-induced
DE   protein PCNT107).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. White Burley;
RX   MEDLINE=91322513; PubMed=1863770;
RA   van der Zaal E.J., Droog F.N.J., Boot C.J.M., Hensgens L.A.M.,
RA   Hoge J.H.C., Schilperoort R.A., Libbenga K.R.;
RT   "Promoters of auxin-induced genes from tobacco can lead to auxin-
RT   inducible and root tip-specific expression.";
RL   Plant Mol. Biol. 16:983-998(1991).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- TISSUE SPECIFICITY: Root tip-specific expression.
CC   -!- INDUCTION: By auxin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X56266; CAA39707.1; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase; Multigene family.
SQ   SEQUENCE   221 AA;  25789 MW;  1488D87B93BC4C60 CRC64;

  Query Match           34.1%;  Score 407.5;  DB 1;  Length 221;
  Best Local Similarity   40.3%;  Pred. No. 8.7e-26;
  Matches   94;  Conservative   29;  Mismatches   87;  Indels   23;  Gaps    4;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||  :: ||  | | | :|::|||  | : ||  || |  || |||: ||:||||||  |
Db          1 MANEEVILLDFWPSMFGMRLRIALAEKEIKYEYKEEDLRNKSPLLLQMNPIHKKIPVLIH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK--------WLTSL 112
                | |||| | ||||:||| : | ::|| : |||| ||||  ||| |        | |
Db         61 NGKPICESIIAVEYIEEVWKDKAPNLLPSDPYDRAQARFWADYIDKKLYDFGRKLWTTK- 119

Qy        113 KSVLATEDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIR 172
                  |: || |  | :  :|||           | ||||:: |||||    : |:
Db        120 -----GEEQEAAKKDFIECLKVLE-------GALGDKPYFGGESFGFVDIALIGYYSWFY 167

Qy        173 VSENMNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKW 225
                  |         | : |    ||:      :|   ||:  |::|| |:|: |:
Db        168 AYETFG--NFSTEAECPKFVAWAKRCMQRESVAKSLPDQPKVLEFVKVLRQKF 218


RESULT 5
GTXA_TOBAC
ID   GTXA_TOBAC     STANDARD;      PRT;   220 AA.
AC   P25317;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Probable glutathione S-transferase parA (EC 2.5.1.18) (Auxin-regulated
DE   protein parA) (STR246C protein).
GN   PARA OR PAR.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf mesophyll;
RX   MEDLINE=90083251; PubMed=2594768;
RA   Takahashi Y., Kuroda H., Tanaka T., Machida Y., Takebe I., Nagata T.;
RT   "Isolation of an auxin-regulated gene cDNA expressed during the
RT   transition from G0 to S phase in tobacco mesophyll protoplasts.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:9279-9283(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NK326;
RX   MEDLINE=95036027; PubMed=7948901;
RA   Froissard D., Gough C., Czernic P., Schneider M., Toppan A., Roby D.,
RA   Marco Y.;
RT   "Structural organization of str 246C and str 246N, plant defense-
RT   related genes from Nicotiana tabacum.";
RL   Plant Mol. Biol. 26:515-521(1994).
RN   [3]
RP   SEQUENCE OF 1-106 FROM N.A.
RX   MEDLINE=91045927; PubMed=2236015;
RA   Takahashi Y., Niwa Y., Machida Y., Nagata T.;
RT   "Location of the cis-acting auxin-responsive region in the promoter
RT   of the par gene from tobacco mesophyll protoplasts.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:8013-8016(1990).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- INDUCTION: By auxin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D90215; BAA14243.1; -.
DR   EMBL; M29274; AAA67894.1; -.
DR   EMBL; X80829; CAA56790.1; -.
DR   EMBL; X80828; CAA56789.1; -.
DR   PIR; A36225; A36225.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase; Multigene family.
SQ   SEQUENCE   220 AA;  25225 MW;  2433BEA9D71D2AC8 CRC64;

  Query Match           33.6%;  Score 402;  DB 1;  Length 220;
  Best Local Similarity   41.2%;  Pred. No. 2.4e-25;
  Matches   94;  Conservative   38;  Mismatches   78;  Indels   18;  Gaps    7;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |   :: ||  | | | :|::||| |||: ||  || |: || |||: ||||||||:  |
Db          1 MESNNVVLLDFWPSSFGMRLRIALALKGIKYEAKEENLSDKSPLLLEMNPVHKKIPILIH 60
```

```
Qy          61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLA--- 117
               | ||||  |:|||||| ::   :|| : |:|: ||||  ||| |  :: : | :
Db          61 NSKAICESLNILEYIDEVW-HDKCPLLPSDPYERSQARFWADYIDKKIYSTGRRVWSGKG 119

Qy         118 TEDDEAKKLHFEQAEEVLEKVE-EVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSEN 176
                : :|||    :  |:|: :| |: |     | ||||| :|||:    | |:    |
Db         120 EDQEEAKK----EFIEILKTLEGELGN-----KTYFGGDNLGFVDVALVPFTSWFYSYET 170

Qy         177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLK 224
                     :  | : | | :||:|     :|   ||   |:  |  :|:||
Db         171 CANFSI--EAECPKLVVWAKTCMESESVSKSLPHPHKIYGF--VLELK 214


RESULT 6
GTXA_ARATH
ID   GTXA_ARATH     STANDARD;      PRT;   224 AA.
AC   P46421;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glutathione S-transferase 103-1A (EC 2.5.1.18).
GN   103-1A OR AT2G29450 OR F16P2.17.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RX   MEDLINE=96194465; PubMed=8624414;
RA   van der Kop D.A.M., Schuyer M., Scheres B., van der Zaal B.J.,
RA   Hooykaas P.J.J.;
RT   "Isolation and characterization of an auxin-inducible glutathione S-
RT   transferase gene of Arabidopsis thaliana.";
RL   Plant Mol. Biol. 30:839-844(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia; TISSUE=Cotyledon, and Hypocotyl;
RA   Watahiki M., Yamamoto K.;
RL   Submitted (XXX-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Sopory S.K.;
RT   "Nucleotide sequence of glutathione S-transferase cDNA from
RT   Arabidopsis thaliana.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- INDUCTION: By auxin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X89216; CAA61504.1; -.
DR   EMBL; U30489; AAA74019.1; -.
DR   EMBL; D44465; BAA07917.1; -.
DR   EMBL; AF144382; AAD34992.1; -.
DR   EMBL; AC004561; AAC95193.1; -.
DR   PIR; S66354; S66354.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
SQ   SEQUENCE   224 AA;  26000 MW;  BA2F5C06B94FFCC4 CRC64;

  Query Match             33.4%;  Score 399;  DB 1;  Length 224;
  Best Local Similarity   41.1%;  Pred. No. 4.3e-25;
  Matches   90;  Conservative   39;  Mismatches   80;  Indels   10;  Gaps    5;

Qy           1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFF 59
               |||: :::||| | |||: ||::|| |||: || ||| |  || |||  ||:|||:||
Db           1 MAEKEEVKLLGIWASPFSRRVEMALKLKGIPYEYVEEILENKSPLLLALNPIHKKVPVLV 60

Qy          60 HGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATE 119
               |  | | || :|:|||||  |  |   ||||: |:|: |||:   :|:: :      :|
Db          61 HNGKTILESHVILEYIDETWPQN--PILPQDPYERSKARFFAKLVDEQIMNVGFISMARA 118

Qy         120 DDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFI--RVSENM 177
               |:: :::  ||  |::  :|    |   || |||| |:||:|   || : |   | | :
Db         119 DEKGREVLAEQVRELIMYLE----KELVGKDYFGGKTVGFLDFVAGSLIPFCLERGWEGI 174

Qy         178 NERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVE 216
                  ::: | |:|    |          ||  :|  |: ||
Db         175 G-LEVITEEKFPEFKRWVRNLEKVEIVKDCVPPREEHVE 212


RESULT 7
GTX1_TOBAC
ID   GTX1_TOBAC     STANDARD;      PRT;   223 AA.
AC   Q03662;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable glutathione S-transferase (EC 2.5.1.18) (Auxin-induced
DE   protein PGNT1/PCNT110).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. White Burley, and cv. Samsun NN; TISSUE=Leaf;
RX   MEDLINE=91322513; PubMed=1863770;
RA   van der Zaal E.J., Droog F.N.J., Boot C.J.M., Hensgens L.A.M.,
RA   Hoge J.H.C., Schilperoort R.A., Libbenga K.R.;
RT   "Promoters of auxin-induced genes from tobacco can lead to auxin-
RT   inducible and root tip-specific expression.";
RL   Plant Mol. Biol. 16:983-998(1991).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- TISSUE SPECIFICITY: Root tip-specific expression.
CC   -!- INDUCTION: By auxin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X56268; CAA39709.1; -.
DR   EMBL; X56264; CAA39705.1; -.
```

```
DR   PIR; S16267; S16267.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase; Multigene family.
SQ   SEQUENCE   223 AA;  25667 MW;  0B29A74FC15869BD CRC64;

  Query Match             32.8%;  Score 392;  DB 1;  Length 223;
  Best Local Similarity   39.0%;  Pred. No. 1.6e-24;
  Matches   90;  Conservative   43;  Mismatches   88;  Indels   10;  Gaps    4;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |||  ::||| |:|||: ||: || :||: || :||  : || |||:|||:|||:||  |
Db          1 MAE--VKLLGFWYSPFSRRVEWALKIKGVKYEYIEEDRDNKSSLLLQSNPIHKKVPVLIH 58

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               | | || :|:|||||  :     ||||:: |||| ||||  ::|||    :|: |   :
Db         59 NGKRIVESMVILEYIDETFEGP--SILPKDPYDRALARFWAKFLDDKVPAVVKTFLRKGE 116

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              :: |       ||| | :: : |:  : | :| ||  || ||       :: | |  :
Db        117 EQEK-----DKEEVCEMLKVLDNELKD-KKFFVGDKFGFADIAANLVAFWLGVFEEASGV 170

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
                |:   |:|    |   :     :|  ||  ::|: | :       ||:|:|
Db        171 VLVTSEKFPNFCKWRGEYINCSQIKESLPPRDELLAFYRSRCQAAAASASA 221


RESULT 8
LGUL_SOYBN
ID   LGUL_SOYBN     STANDARD;      PRT;   219 AA.
AC   P46417;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lactoylglutathione lyase (EC 4.4.1.5) (Methylglyoxalase)
DE   (Aldoketomutase) (Glyoxalase I).
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Koellner B., Finkelnburg B., Mayerbacher R., Paulus C.,
RA   Springer B.;
RL   Submitted (SEP-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the conversion of hemimercaptal, formed from
CC       methylglyoxal and glutathione, to S-lactoylglutathione.
CC   -!- CATALYTIC ACTIVITY: (R)-S-lactoylglutathione = glutathione +
CC       methylglyoxal.
CC   -!- PATHWAY: Glyoxal pathway.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X68819; CAA48717.1; -.
DR   PIR; S47177; S47177.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Lyase.
SQ   SEQUENCE   219 AA;  25901 MW;  6A2E46B759476A8C CRC64;

  Query Match             31.6%;  Score 378;  DB 1;  Length 219;
  Best Local Similarity   39.7%;  Pred. No. 2e-23;
  Matches   92;  Conservative   28;  Mismatches   78;  Indels   34;  Gaps    7;

Qy          8 LLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVICE 67
              ||  | | : :| :|||  ||: ||  || |  :| |||: ||:||||||  |  | |||
Db          7 LLDTWASMYGMRARIALAEKGVRYEYKEENLMNRSPLLLQMNPIHKKIPVLIHNGKPICE 66

Qy         68 SAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK--------WLTSLKSVLATE 119
              |||||:|||||| |:   ::| : | |: ||||| ||| |         ||:        |
Db         67 SAIIVQYIDEVW-NDKSPLMPSDPYKRSQARFWVDYIDKKIYDTWKKMWLSK-----GEE 120

Qy        120 DDEAKK---LHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSEN 176
               :| ||      |:| || |              | ::| || ||||:   :| |:    |
Db        121 HEEGKKELISIFKQLEETL----------TDKPFYGDDTFGFVDLCLITFSSWFYTYET 169

Qy        177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEK----LVEFAKILQLK 224
                  |:  | : | |  | :        |   ||: :|    :||  | |: |
Db        170 YGNFKM--EEECPKLMAWVKRCMERETVSNTLPDAKKVYGLIVELQKTLESK 219


RESULT 9
GTX2_TOBAC
ID   GTX2_TOBAC     STANDARD;      PRT;   223 AA.
AC   Q03663;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable glutathione S-transferase (EC 2.5.1.18) (Auxin-induced
DE   protein PGNT35/PCNT111).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. White Burley, and cv. Samsun NN; TISSUE=Leaf;
RX   MEDLINE=91322513; PubMed=1863770;
RA   van der Zaal E.J., Droog F.N.J., Boot C.J.M., Hensgens L.A.M.,
RA   Hoge J.H.C., Schilperoort R.A., Libbenga K.R.;
RT   "Promoters of auxin-induced genes from tobacco can lead to auxin-
RT   inducible and root tip-specific expression.";
RL   Plant Mol. Biol. 16:983-998(1991).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- TISSUE SPECIFICITY: Root tip-specific expression.
CC   -!- INDUCTION: By auxin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X56269; CAA39710.1; -.
DR   EMBL; X56265; CAA39706.1; -.
DR   PIR; S16268; S16268.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase; Multigene family.
SQ   SEQUENCE   223 AA;  25841 MW;  ACD523DB2A003A83 CRC64;

  Query Match             31.3%;  Score 373.5;  DB 1;  Length 223;
  Best Local Similarity   37.9%;  Pred. No. 4.8e-23;
  Matches   88;  Conservative   46;  Mismatches   85;  Indels   13;  Gaps    6;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |||  ::||| |:|||: ||: || :||: || :||  : || |||:||||:||:||  |
```

```
Db          1 MAE--VKLLGFWYSPFSHRVEWALKIKGVKYEYIEEDRDNKSSLLLQSNPVYKKVPVLIH 58

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               | | || ||:|||| :     ||||:: |||| |||  ::||| : :: :    :
Db         59 NGKPIVESMIILEYIDETFEGP--SILPKDPYDRALARFWAKFLDDK-VAAVVNTFFRKG 115

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              :| :|      ||| | :: : |:  : | :| ||  || ||        :: | |
Db        116 EEQEK----GKEEVYEMLKVLDNELKD-KKFFAGDKFGFADIAANLVGFWLGVFEEGYGD 170

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAK 232
                |:   |:|  : | : :     |   ||  ::|: |    : :: |  |::
Db        171 VLVKSEKFPNFSKWRDEYINCSQVNESLPPRDELLAF---FRARFQAVVASR 219


RESULT 10
GTX1_SOLTU
ID   GTX1_SOLTU     STANDARD;      PRT;   217 AA.
AC   P32111;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable glutathione S-transferase (EC 2.5.1.18) (Pathogenesis-related
DE   protein 1).
GN   PRP1.
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92063197; PubMed=2132026;
RA   Taylor J.L., Fritzemeier K.H., Haeuser I., Kombrink E., Rohwer F.,
RA   Schroeder M., Strittmatter G., Hahlbrock K.;
RT   "Structural analysis and activation by fungal infection of a gene
RT   encoding a pathogenesis-related protein in potato.";
RL   Mol. Plant Microbe Interact. 3:72-77(1990).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- INDUCTION: By fungal infection.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J03679; AAA68430.1; -.
DR   PIR; T07595; T07595.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Plant defense; Transferase; Pathogenesis-related protein.
SQ   SEQUENCE   217 AA;  25056 MW;  7B0DBBE216685B4E CRC64;

  Query Match             31.1%;  Score 372;  DB 1;  Length 217;
  Best Local Similarity   41.4%;  Pred. No. 6.1e-23;
  Matches   82;  Conservative   35;  Mismatches   71;  Indels   10;  Gaps    4;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |||  ::|||  :|||: ||: || :||: || :|| |  || |||:|||:|||||  |
Db          1 MAE--VKLLGLRYSPFSHRVEWALKIKGVKYEFIEEDLQNKSPLLLQSNPIHKKIPVLIH 58

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
                | |||| :|:||||| :     ||||:: |||| ||||  |::||     ||  :  :
Db         59 NGKCICESMVILEYIDEAFEGP--SILPKDPYDRALARFWAKYVEDKGAAVWKSFFSKGE 116

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              :: |    |:| |:|: ::  |    : |  | ||  || ||       :: : | ::
```

```
Db        117 EQEKAK--EEAYEMLKILDNEF----KDKKCFVGDKFGFADIVANGAALYLGILEEVSGI 170

Qy        181 KLLDETKYPGLTLWAETF 198
                |    |:|    | : :
Db        171 VLATSEKFPNFCAWRDEY 188


RESULT 11
GTX3_TOBAC
ID   GTX3_TOBAC     STANDARD;      PRT;   223 AA.
AC   Q03664;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Probable glutathione S-transferase (EC 2.5.1.18) (Auxin-induced
DE   protein PCNT103).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. White Burley;
RX   MEDLINE=91322513; PubMed=1863770;
RA   van der Zaal E.J., Droog F.N.J., Boot C.J.M., Hensgens L.A.M.,
RA   Hoge J.H.C., Schilperoort R.A., Libbenga K.R.;
RT   "Promoters of auxin-induced genes from tobacco can lead to auxin-
RT   inducible and root tip-specific expression.";
RL   Plant Mol. Biol. 16:983-998(1991).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- TISSUE SPECIFICITY: Root tip-specific expression.
CC   -!- INDUCTION: By auxin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X56263; CAA39704.1; -.
DR   PIR; S16269; S16269.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase; Multigene family.
SQ   SEQUENCE   223 AA;  25747 MW;  22CD25BB5A05D214 CRC64;

  Query Match             31.1%;  Score 371.5;  DB 1;  Length 223;
  Best Local Similarity   39.1%;  Pred. No. 6.9e-23;
  Matches   91;  Conservative   42;  Mismatches   89;  Indels   11;  Gaps    6;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |||  ::||| |:|||  ||: || |||: || :||  : || |||:|||||||:||  |
Db          1 MAE--VKLLGFWYSPFTHRVEWALKLKGVKYEYIEEDRDNKSSLLLQSNPVHKKVPVLIH 58

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
                | | || :|:||||| :     ||||:: |||| ||||  :: || : :: :    :
Db         59 NGKPIVESMVILEYIDETFEGP--SILPKDPYDRALARFWSKFLGDK-VAAVVNTFFRKG 115

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              :| :|      ||| | :: : |:  : | :| ||  || ||       :: | |
Db        116 EEQEK----GKEEVYEMLKVLDNELKD-KKFFVGDKFGFADIAANLVGFWLGVFEEGYGV 170

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKI-LQLKWAAAAAAK 232
                |:   |:|  : | : :     ||  ||  ::|: | :    |    |: :| |
Db        171 VLVTSEKFPNFSRWRDEYINCSQVKESLPSRDELLAFFRARFQAVVASISAPK 223
```

```
RESULT 12
GTX1_NICPL
ID   GTX1_NICPL     STANDARD;      PRT;   219 AA.
AC   P50471;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Probable glutathione S-transferase MSR-1 (EC 2.5.1.18) (Auxin-
DE   regulated protein MSR-1).
GN   MSR-1.
OS   Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4092;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92361262; PubMed=1498603;
RA   Dominov J.A., Stenzler L., Lee S., Schwarz J.J., Leisner S.,
RA   Howell S.H.;
RT   "Cytokinins and auxins control the expression of a gene in Nicotiana
RT   plumbaginifolia cells by feedback regulation.";
RL   Plant Cell 4:451-461(1992).
CC   -!- FUNCTION: May play an important role in hormonal and growth
CC       regulatory responses.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- INDUCTION: By auxin and cytokinin.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S44036; AAB47712.2; -.
DR   PIR; JQ1606; JQ1606.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase; Multigene family.
SQ   SEQUENCE   219 AA;  25269 MW;  DCE165B0FB33A759 CRC64;

  Query Match             30.3%;  Score 361.5;  DB 1;  Length 219;
  Best Local Similarity   39.0%;  Pred. No. 4.3e-22;
  Matches   89;  Conservative   35;  Mismatches   85;  Indels   19;  Gaps    7;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |   :: ||    | | :|::||| |||: ||  || |: || |||: |||||||:  |
Db          1 MESNNVVLLDFSGSSFGMRLRIALALKGIKYEAKEENLSDKSPLLLEMNPVHKKIPILIH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               | ||||  |:|||||| :    :|| : | |: |||| :|||:|  :: : | : :
Db         61 NGKPICESLNILEYIDEVW-HEKCPLLPSDPYQRSQARFWANYIDNKIYSTGRRVWSGKG 119

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEG----KAYFGGDTIGFVDIGFGSFLSFIRVSEN 176
                    |  ||  ::  |:| |  ||    | ||||| :||||:    | |:    |
Db        120 --------EDQEEAKKEFIEIF-KTLEGELGNKTYFGGDNLGFVDVALVPFTSWFYSYET 170

Qy        177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLK 224
                  :  | :   | :|      :   |  ||   |: :|  :|:||
Db        171 CANFSI--EAECRKLVVWQNCMENERVSKS-LPHPHKIYDF--VLELK 213


RESULT 13
GTXC_ORYSA
ID   GTXC_ORYSA     STANDARD;      PRT;   254 AA.
AC   Q06398;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Probable glutathione S-transferase (EC 2.5.1.18) (28 kDa cold-induced
DE   protein).
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare; TISSUE=Seedling;
RA   Binh L.T., Oono K.;
RT   "Molecular cloning and characterization of genes related to
RT   chilling tolerance in rice.";
RL   Plant Physiol. 99:1146-1150(1992).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- TISSUE SPECIFICITY: Seedling shoots and roots.
CC   -!- SIMILARITY: Belongs to the GST superfamily. HSP26 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D10861; BAA01632.1; -.
DR   Gramene; Q06398; -.
DR   InterPro; IPR004045; GST_Nterm.
DR   Pfam; PF02798; GST_N; 1.
KW   Transferase.
FT   DOMAIN      121    129       GLY-RICH.
FT   DOMAIN      158    168       ARG-RICH.
SQ   SEQUENCE   254 AA;  28241 MW;  8D3FD11AAA782A8F CRC64;

  Query Match             17.0%;  Score 203.5;  DB 1;  Length 254;
  Best Local Similarity   44.7%;  Pred. No. 2.6e-09;
  Matches   42;  Conservative   17;  Mismatches   30;  Indels    5;  Gaps    2;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHG--- 61
              :|:|||| | ||:|:||:: |||| | || ||| |  ||:||| |||:||| :||  |
Db          6 ELKLLGVWSSPYAIRVRVVLNLKSLPYEYVEENLGDKSDLLLASNPVHKSVPVLLHAGRR 65

Qy         62 DKVICESAIIVEYIDEVWSNNALSILPQNAYDRA 95
              ::|    |:    :    :    |::| : |:||
Db         66 ERVAGHRAVHRRGL--AGARRGRSVMPSDPYERA 97


RESULT 14
GTO1_PIG
ID   GTO1_PIG       STANDARD;      PRT;   241 AA.
AC   Q9N1F5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutathione transferase omega 1 (EC 2.5.1.18) (GSTO 1-1) (Glutathione-
DE   dependent dehydroascorbate reductase).
GN   GSTO1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 5-31; 58-83; 101-110; 115-147;
RP   149-160; 162-190; 201-228 AND 235-241, SUBUNIT, SUBCELLULAR LOCATION,
RP   TISSUE SPECIFICITY, AND MASS SPECTROMETRY.
RC   TISSUE=Liver;
RX   MEDLINE=21378194; PubMed=11485575;
RA   Rouimi P., Anglade P., Benzekri A., Costet P., Debrauwer L.,
RA   Pineau T., Tulliez J.;
RT   "Purification and characterization of a glutathione S-transferase
RT   Omega in pig: evidence for two distinct organ-specific transcripts.";
```

```
RL   Biochem. J. 358:257-262(2001).
CC   -!- FUNCTION: Exhibits glutathione-dependent thiol transferase and
CC       dehydroascorbate reductase activities (By similarity).
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Most abundant in the liver and skeletal
CC       muscle; also expressed in heart, diaphragm, colon, thymus, kidney,
CC       lung, ovaries, spleen, intestine and pancreas.
CC   -!- MASS SPECTROMETRY: MW=27328; MW_ERR=3; METHOD=Electrospray.
CC   -!- SIMILARITY: Belongs to the GST superfamily. Omega family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF188838; AAF71994.2; -.
DR   HSSP; P78417; 1EEM.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR005442; GST_omega.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01625; GSTRNSFRASEO.
KW   Transferase.
FT   ACT_SITE     32     32       BY SIMILARITY.
FT   CONFLICT    139    139       C -> Y (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    158    158       T -> Q (IN REF. 1; AA SEQUENCE).
SQ   SEQUENCE   241 AA;  27419 MW;  AA50EE81C70433A6 CRC64;

  Query Match          16.2%;  Score 193;  DB 1;  Length 241;
  Best Local Similarity   30.0%;  Pred. No. 1.7e-08;
  Matches   64;  Conservative   33;  Mismatches  104;  Indels   12;  Gaps    8;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGD-KV 64
              :|:    | ||| |  : || ||: ::|:   |  | |   : ||    :||  :   ::
Db         24 IRVYSMRFCPFAQRTLLVLNAKGIRHQVININLKNKPEWFFQKNP-SGLVPVLENSQGQL 82

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              | ||||  ||:|| :    |  || : |::|  :  |  :  |    |   :  |::
Db         83 IYESAITCEYLDEAYPGKKL--LPDDPYEKACQKM-VFELSSKVPPLLIRFIRRENEADC 139

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
                  |:  :   |:|||  |  :   |||| ::  :|     :   :   | :||   :|
Db        140 SGLKEELRKEFSKLEEVLTK--KKTTYFGGSSLSMIDYLIWPWFERLEALE-LNE--CID 194

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEF 217
              |  | | ||       ||||  |  |   | |
Db        195 HT--PKLKLWMAAMMKDPAVSALHIEPRDLRAF 225


RESULT 15
GTO1_HUMAN
ID   GTO1_HUMAN     STANDARD;      PRT;   241 AA.
AC   P78417;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Glutathione transferase omega 1 (EC 2.5.1.18) (GSTO 1-1).
GN   GSTO1 OR GSTTLP28.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Kodym R., Story M.D.;
RT   "Cloning of the human homolog to a mouse protein, differentially
RT   expressed in lymphoma cells with different susceptibility to radiation
RT   induced apoptosis.";
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RC   TISSUE=Fetal;
RX   MEDLINE=20387379; PubMed=10783391;
RA   Board P.G., Coggan M., Chelvanayagam G., Easteal S., Jermiin L.S.,
RA   Schulte G.K., Danley D.E., Hoth L.R., Griffor M.C., Kamath A.V.,
RA   Rosner M.H., Chrunyk B.A., Perregaux D.E., Gabel C.A., Geoghegan K.F.,
RA   Pandit J.;
RT   "Identification, characterization, and crystal structure of the Omega
RT   class glutathione transferases.";
RL   J. Biol. Chem. 275:24798-24806(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 12-21; 58-65; 133-139 AND 149-156, AND MASS SPECTROMETRY.
RC   TISSUE=Liver;
RX   MEDLINE=21088920; PubMed=11271497;
RA   Hubbard M.J., McHugh N.J.;
RT   "Human ERp29: isolation, primary structural characterisation and two-
RT   dimensional gel mapping.";
RL   Electrophoresis 21:3785-3796(2000).
RN   [5]
RP   VARIANTS ASP-140 AND GLU-155 DEL.
RX   MEDLINE=22506234; PubMed=12618591;
RA   Whitbread A.K., Tetlow N., Eyre H.J., Sutherland G.R., Board P.G.;
RT   "Characterization of the human Omega class glutathione transferase
RT   genes and associated polymorphisms.";
RL   Pharmacogenetics 13:131-144(2003).
CC   -!- FUNCTION: Exhibits glutathione-dependent thiol transferase and
CC       dehydroascorbate reductase activities.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Ubiquitous. Highest expression in liver,
CC       skeletal muscle and heart. Lowest expression in brain, placenta
CC       and lung.
CC   -!- SIMILARITY: Belongs to the GST superfamily. Omega family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U90313; AAB70109.1; -.
DR   EMBL; AF212303; AAF73376.1; -.
```

```
DR   EMBL; BC000127; AAH00127.1; -.
DR   PDB; 1EEM; 11-AUG-00.
DR   Genew; HGNC:13312; GSTO1.
DR   MIM; 605482; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0004364; F:glutathione transferase activity; NAS.
DR   GO; GO:0016656; F:monodehydroascorbate reductase (NADH) activity; NAS.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR005442; GST_omega.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01625; GSTRNSFRASEO.
KW   Transferase; Polymorphism; 3D-structure.
FT   ACT_SITE     32     32
FT   VARIANT     140    140       A -> D (in allele GSTO1*C).
FT                                /FTId=VAR_016811.
FT   VARIANT     155    155       Missing (in allele GSTO1*B).
FT                                /FTId=VAR_016813.
FT   STRAND        9      9
FT   TURN         11     12
FT   TURN         21     22
FT   STRAND       24     28
FT   TURN         30     31
FT   HELIX        33     44
FT   TURN         45     46
FT   STRAND       49     54
FT   TURN         56     57
FT   HELIX        61     65
FT   TURN         66     66
FT   TURN         68     69
FT   STRAND       74     77
FT   TURN         78     79
FT   STRAND       82     84
FT   HELIX        86     96
FT   HELIX       107    120
FT   TURN        121    122
FT   HELIX       123    130
FT   TURN        131    132
FT   HELIX       136    160
FT   STRAND      164    164
FT   TURN        165    166
FT   STRAND      167    167
FT   HELIX       172    184
FT   TURN        185    188
FT   HELIX       190    192
FT   HELIX       197    207
FT   TURN        208    208
FT   HELIX       210    215
FT   HELIX       219    231
FT   TURN        232    232
FT   TURN        234    235
FT   HELIX       236    238
FT   TURN        239    241
SQ   SEQUENCE   241 AA;  27566 MW;  9134ABA265F5C87E CRC64;

  Query Match            16.1%;  Score 192;  DB 1;  Length 241;
  Best Local Similarity  30.0%;  Pred. No. 2e-08;
  Matches   62;  Conservative   38;  Mismatches   95;  Indels   12;  Gaps    8;

Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGD 62
              |  :|:    | ||| | :: |  ||: :||:   |  | |   | ||    :||  :
Db         21 EGSIRIYSMRFCPFAERTRLVLKAKGIRHEVININLKNKPEWFFKKNP-FGLVPVLENSQ 79

Qy         63 -KVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDD 121
               ::| ||||  ||:|| :    |  || : |::|  :  :  :  |  : : | : :::
Db         80 GQLIYESAITCEYLDEAYPGKKL--LPDDPYEKACQKM-ILELFSKVPSLVGSFIRSQNK 136

Qy        122 EAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERK 181
              |      |:  :   |:|||    ::   :|||::|  :|     :  | |: | |   :
Db        137 EDYAGLKEEFRKEFTKLEEVLT--NKKTTFFGGNSISMIDYLIWPW--FERL-EAMKLNE 191

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLL 208

              :| |  | | ||      || |  ||
Db        192 CVDHT--PKLKLWMAAMKEDPTVSALL 216


Search completed: October  8, 2004, 09:22:10
Job time : 26 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 8, 2004, 09:16:45 ; Search time 39 Seconds
(without alignments)
572.216 Million cell updates/sec

Title: US-10-088-945A-10
Perfect score: 1195
Sequence: 1 MAERDLRLLGAWFSPFALRV.....KLVEFAKILQLKWAAAAAAK 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	596	49.9	234	2	F96620	hypothetical prote
2	558	46.7	233	2	D96620	protein T30E16.25
3	520	43.5	470	2	H86397	protein T7N9.20 [i
4	491.5	41.1	234	2	G96721	probable glutathio
5	486.5	40.7	229	2	F96721	probable glutathio
6	429.5	35.9	225	2	A33654	heat shock protein
7	426	35.6	224	2	A84697	probable glutathio
8	412.5	34.5	241	2	S22457	Bronze-2 protein -
9	411	34.4	230	2	T10825	auxin-induced prot
10	410.5	34.4	221	2	S19185	parC protein - com
11	407.5	34.1	221	2	S16636	auxin-induced prot
12	403	33.7	161	2	T09807	probable glutathio
13	402	33.6	220	2	A36225	auxin-regulated pr
14	399	33.4	224	2	S66354	glutathione transf
15	397.5	33.3	225	2	G84696	probable glutathio
16	396.5	33.2	218	2	T09781	glutathione transf
17	394	33.0	219	2	T06239	probable glutathio
18	392	32.8	223	2	S16267	auxin-induced prot
19	390	32.6	216	2	T07156	probable glutathio
20	383.5	32.1	225	2	H84696	probable glutathio
21	378	31.6	219	2	S47177	lactoylglutathione
22	377	31.5	224	2	F84696	probable glutathio
23	376	31.5	220	2	A86308	probable glutathio
24	373.5	31.3	223	2	S16268	auxin-induced prot
25	372.5	31.2	227	2	B84696	probable glutathio
26	372	31.1	217	2	T07595	glutathione transf
27	371.5	31.1	223	2	S16269	auxin-induced prot
28	371	31.0	220	2	C96812	protein F3F9.14 [i
29	371	31.0	223	2	D84696	probable glutathio
30	371	31.0	224	2	T12332	glutathione transf
31	365	30.5	224	2	T04358	glutathione transf
32	364	30.5	219	2	T51607	glutathione transf
33	361.5	30.3	219	2	JQ1606	multiple stimulus
34	361	30.2	224	2	A96577	probable glutathio
35	359	30.0	218	2	G86307	probable glutathio
36	357.5	29.9	219	2	S19182	gene C-7 protein -
37	353.5	29.6	221	2	H86307	probable glutathio
38	318	26.6	232	2	A96775	probable glutathio
39	317	26.5	227	2	T47416	glutathione transf
40	307	25.7	231	2	T02765	glutathione transf
41	204	17.1	102	2	F86237	protein F14N23.25
42	180.5	15.1	257	2	S33314	beta-etherase ligF
43	170.5	14.3	266	2	T20806	hypothetical prote
44	164	13.7	213	2	T06333	probable glutathio
45	164	13.7	220	2	S33628	glutathione transf

ALIGNMENTS

RESULT 1
F96620
hypothetical protein F23H11.1 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: F96620
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J. Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Aansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96620
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <STO>
A;Cross-references: GB:AE005173; NID:g5080803; PIDN:AAD39312.1; GSPDB:GN00141
C;Genetics:
A;Gene: F23H11.1
A;Map position: 1
C;Superfamily: auxin-induced protein

Query Match 49.9%; Score 596; DB 2; Length 234;
Best Local Similarity 49.6%; Pred. No. 3.6e-41;
Matches 115; Conservative 46; Mismatches 69; Indels 2; Gaps 2;

```
Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | |: :::||| |:||:|:| :||| || :||: ||| |   ||||||||||||||:||
Db          1 MGEKEEVKLLGVWYSPYAIRPKIALRLKSVDYDYVEENLFGSKSELLLKSNPVHKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
               | :| | ||  ||||||| |:::| |||| : |||| ||||  ::|:|| :|:    |
Db         61 LHNNKPIVESLNIVEYIDETWNSSAPSILPSHPYDRALARFWSDFVDNKWFPALRMAAIT 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              : ::||    |: || | ::|: |   |:|| :|||: |||:|| ||||:  ::  |
Db        121 KSEDAKAKAMEEVEEGLLQLEDAFVSISKGKPFFGGEAIGFMDICFGSFVVLLKAREKFK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              |||||:| | |  ||: | :|  || : || ||: || : |:::  :||:
Db        181 AEKLLDESKTPSLCKWADRFLSDETVKNVAPEIEKVAEFLQELEVRAQSAAS 232
```

RESULT 2
D96620
protein T30E16.25 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: D96620
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96620
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <STO>
A;Cross-references: GB:AE005173; NID:g8778750; PIDN:AAF79758.1; GSPDB:GN00141
C;Genetics:
A;Gene: T30E16.25
A;Map position: 1
C;Superfamily: auxin-induced protein

Query Match 46.7%; Score 558; DB 2; Length 233;
Best Local Similarity 48.4%; Pred. No. 4.4e-38;
Matches 106; Conservative 42; Mismatches 69; Indels 2; Gaps 2;

```
Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | || :::||| |:|| :| :||| || :||: ||| |   ||||||||||: ||:||
Db          1 MGEREEVKLLGTWYSPVVIRAKIALRLKSVDYDYVEEDLFGSKSELLLKSNPIFKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
               | | :| |  ||||||| |:::  |||| : |||| ||||  ::||||| :| : :
Db         61 IHNTKPVCVSLNIVEYIDETWNSSGSSILPSHPYDRALARFWSVFVDDKWLPTLMAAVVA 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              : :|||    |: || | ::|  |   |:||::|||:||||:||  ||||  ::  | :
Db        121 KSEEAKAKGMEEVEEGLLQLEAAFIALSKGKSFFGGETIGFIDICLGSFLVLLKAREKLK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEF 217
              |:||| | | |  ||  | ::  || ::|: :|: :|
Db        181 NEKILDELKTPSLYRWANQFLSNEMVKNVVPDIDKVAKF 219
```

RESULT 3
H86397
protein T7N9.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: H86397
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86397
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-470 <STO>
A;Cross-references: GB:AE005172; NID:g8778860; PIDN:AAF79859.1; GSPDB:GN00141
C;Genetics:
A;Gene: T7N9.20
A;Map position: 1

Query Match 43.5%; Score 520; DB 2; Length 470;
Best Local Similarity 47.7%; Pred. No. 1.3e-34;
Matches 106; Conservative 39; Mismatches 73; Indels 4; Gaps 3;

```
Qy          1 MAERD-LRLLGAWFSPFALRVQIALNLKGLDYEVVEE--TLNPKSELLLKSNPVHKKIPV 57
              ||: | ::|:|:| ||::|| ::||:|| : || ::|    |  |||||||||:|||:||
Db          1 MAQNDTVKLIGSWSSPYSLRARVALHLKSVKYEYLDEPDVLKEKSELLLKSNPIHKKVPV 60

Qy         58 FFHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLA 117
                 |||  | ||  :|:|:|| |  :  |||| :|||||:||||  |||||   :: :|:
Db         61 LLHGDLSISESLNVVQYVDEAWP-SVPSILPSDAYDRASARFWAQYIDDKCFAAVDAVVG 119

Qy        118 TEDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENM 177
               :||| |     :  | |  :|| | | |:|  :|||:|||::||   : |  | | |
Db        120 AKDDEGKMAAVGKLMECLAILEETFQKSSKGLGFFGGETIGYLDIACSALLGPISVIEAF 179

Qy        178 NERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              :  | | :   |||  ||| | |  |||  :|  |::| |||
Db        180 SGVKFLRQETTPGLIKWAERFRAHEAVKPYMPTVEEVVAFAK 221
```

RESULT 4
G96721
probable glutathione transferase T17F3.4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: G96721
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96721
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <STO>
A;Cross-references: GB:AE005173; NID:g6358801; PIDN:AAF07381.1; GSPDB:GN00141
C;Genetics:
A;Gene: T17F3.4
A;Map position: 1
C;Superfamily: auxin-induced protein

Query Match 41.1%; Score 491.5; DB 2; Length 234;
Best Local Similarity 46.0%; Pred. No. 1.1e-32;
Matches 99; Conservative 36; Mismatches 79; Indels 1; Gaps 1;

```
Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              ::||||| ||| || :|||||| : || :||      || :|  |||||:||:  ||:| |
Db         13 VKLLGAWPSPFVLRTRIALNLKNVAYEYLEEEDTLSSESVLNYNPVHKQIPILIHGNKPI 72

Qy         66 CESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAKK 125
               ||  || |:|| | :    ||| : :||| ||||  |||:   ||:  |   : :|
Db         73 RESLNIVMYVDETWLSGP-PILPSDPFDRAVARFWDVYIDEHCFTSINGVAVAKGEENIN 131

Qy        126 LHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLDE 185
                : |: :  :|| | :||:|: :|||: |||:|||||| |  : | |      | :
Db        132 AAIAKLEQCMALLEETFQECSKGRGFFGGENIGFIDIGFGSMLGPLTVLEKFTGVKFIHP 191

Qy        186 TKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKI 220
              |||  ||: | |  ||| ::|: ||||:||::
Db        192 ENTPGLFHWADRFYAHEAVKPVMPDIEKLVQFARL 226
```

RESULT 5
F96721
probable glutathione transferase T17F3.5 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: F96721
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96721
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-229 <STO>
A;Cross-references: GB:AE005173; NID:g6358802; PIDN:AAF07382.1; GSPDB:GN00141
C;Genetics:
A;Gene: T17F3.5
A;Map position: 1
C;Superfamily: auxin-induced protein

Query Match 40.7%; Score 486.5; DB 2; Length 229;
Best Local Similarity 47.2%; Pred. No. 2.8e-32;
Matches 102; Conservative 33; Mismatches 78; Indels 3; Gaps 2;

```
Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEET--LNPKSELLLKSNPVHKKIPVFFHGDK 63
              ::|:| | ||||:| |:||:|| :::| |||| | ||:|||:||||:|||:||  |||
Db         10 VKLIGTWASPFAIRAQVALHLKSVEHEYVEETDVLKGKSDLLIKSNPIHKKVPVLIHGDV 69

Qy         64 VICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEA 123
               ||||  ||:|:|| | :: |||||    :|| ||||  ::| |   |: :|   :|| |
Db         70 SICESLNIVQYVDESWPSD-LSILPTLPSERAFARFWAHFVDGKLFESIDAVAGAKDDAA 128

Qy        124 KKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLL 183
              :       | |  :|| | | |:|  :|||  ||||||  |: :  | | |  :  | |
Db        129 RMTLAGNLMENLAALEEAFQKSSKGGDFFGGGNIGFVDITVGAIVGPISVIEAFSGVKFL 188

Qy        184 DETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
                  |||  ||| | |  |||  :|   : :||||
Db        189 RPDTTPGLIQWAEKFRAHEAVKPYMPTVAEFIEFAK 224
```

RESULT 6
A33654
heat shock protein 26A - soybean
C;Species: Glycine max (soybean)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 17-Nov-2000
C;Accession: A33654; A28662
R;Czarnecka, E.; Nagao, R.T.; Key, J.L.; Gurley, W.B.
Mol. Cell. Biol. 8, 1113-1122, 1988
A;Title: Characterization of Gmhsp26-A, a stress gene encoding a divergent heat shock pr
A;Reference number: A33654; MUID:88216585; PMID:2835661
A;Accession: A33654
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-225 <CZA>
A;Cross-references: GB:M20363; NID:g169980; PIDN:AAA33973.1; PID:g169981
R;Hagen, G.; Uhrhammer, N.; Guilfoyle, T.J.
J. Biol. Chem. 263, 6442-6446, 1988
A;Title: Regulation of expression of an auxin-induced soybean sequence by cadmium.
A;Reference number: A28662; MUID:88198196; PMID:3360788
A;Accession: A28662
A;Molecule type: mRNA
A;Residues: 86-225 <HAG>
C;Genetics:
A;Introns: 107/3
C;Superfamily: auxin-induced protein

Query Match 35.9%; Score 429.5; DB 2; Length 225;
Best Local Similarity 40.7%; Pred. No. 1.2e-27;
Matches 94; Conservative 46; Mismatches 80; Indels 11; Gaps 5;

```
Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGD 62
              : |::|||   |||  |||||| |||::|: :|| |  ||:|||| ||||||:||| | :
Db          5 QEDVKLLGIVGSPFVCRVQIALKLKGVEYKFLEENLGNKSDLLLKYNPVHKKVPVFVHNE 64

Qy         63 KVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDE 122
              : | || :|||||| | ||   ||| : | || |||| :|||| : ::   : | |::
Db         65 QPIAESLVIVEYIDETWKNN--PILPSDPYQRALARFWSKFIDDKIVGAVSKSVFTVDEK 122

Qy        123 AKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSF-IRVSENMNERK 181
               :: : |:  | |: :|       : | :|||:  | |||    |::| | : : :   :
Db        123 EREKNVEETYEALQFLENEL----KDKKFFGGEEFGLVDIA-AVFIAFWIPIFQEIAGLQ 177

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAK 232
              |    ]:| |  |:: |   | |  :||  : |  : |    :: : :|:|
Db        178 LFTSEKFPILYKWSQEFLNHPFVHEVLPPRDPLFAYFK---ARYESLSASK 225
```

RESULT 7
A84697
probable glutathione S-transferase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: A84697
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84697
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-224 <STO>
A;Cross-references: GB:AE002093; NID:g3980386; PIDN:AAC95189.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g29490
A;Map position: 2
C;Superfamily: auxin-induced protein

Query Match 35.6%; Score 426; DB 2; Length 224;
Best Local Similarity 42.1%; Pred. No. 2.2e-27;
Matches 93; Conservative 48; Mismatches 66; Indels 14; Gaps 7;

```
Qy          1 MAERD--LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVF 58
              |||::  ::||| | |||: ||::|| ||| || :|| |  |: |||: ||:|||:||
Db          1 MAEKEESVKLLGFWASPFSRRVEMALKLKGVPYEYLEEDLPNKTPLLLELNPLHKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLT-SLKSVLA 117
               | ||:: || :|:|||: | |:   ||||: |::| ||||  :|||: ||   :|::
Db         61 VHNDKILLESHLILEYIDQTWKNS--PILPQDPYEKAMARFWAKFIDDQILTLGFRSLVK 118

Qy        118 TEDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSF--IRVSE 175
               |  : :::  |:  ]:|  :|    |    || :||| ||||:|:  || : |   |: :
Db        119 AE--KGREVAIEETRELLMFLE----KEVTGKDFFGGKTIGFLDMIAGSMIPFCLARLWK 172

Qy        176 NMNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVE 216
               :    :: | |:| |  | :      ||:| :|  || :|
Db        173 GIG-IDMIPEEKFPELNRWIKNLEEVEAVRGCIPPREKQIE 212
```

RESULT 8
S22457
Bronze-2 protein - maize
C;Species: Zea mays (maize)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
C;Accession: S22457; JQ0987; S58414
R;Schmitz, G.; Theres, K.
Mol. Gen. Genet. 233, 269-277, 1992
A;Title: Structural and functional analysis of the Bz2 locus of Zea mays: characterizati
A;Reference number: S22457; MUID:92293125; PMID:1376405
A;Accession: S22457
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-241 <SCH>
R;Nash, J.; Luehrsen, K.R.; Walbot, V.

```
Plant Cell 2, 1039-1049, 1990
A;Title: Bronze-2 gene of maize: reconstruction of a wild-type allele and analysis of tr
A;Reference number: JQ0987; MUID:93005645; PMID:1967051
A;Accession: JQ0987
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-195,198-233,'VQ',234-241 <NAS>
R;Marrs, K.A.; Alfenito, M.R.; Lloyd, A.M.; Walbot, V.
Nature 375, 397-400, 1995
A;Title: A glutathione S-transferase involved in vacuolar transfer encoded by the maize
A;Reference number: S58414; MUID:95281051; PMID:7760932
A;Contents: annotation
C;Genetics:
A;Gene: Bz2
A;Introns: 114/3
C;Function:
A;Description: required for production of anthocyanin pigment
A;Pathway: anthocyanin biosynthesis
C;Superfamily: auxin-induced protein

  Query Match             34.5%;  Score 412.5;  DB 2;  Length 241;
  Best Local Similarity   39.8%;  Pred. No. 3.1e-26;
  Matches   94;  Conservative   47;  Mismatches   84;  Indels   11;  Gaps    6;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNP-KSELLLKSNPVHKKIPVFF 59
              |     :|:||   |||   | ::||:|:|: ||:::| | | ||: || :|||: ||||
Db          1 MTAGTMRVLGGEVSPFTARARLALDLRGVAYELLDEPLGPKKSDRLLAANPVYGKIPVLL 60

Qy         60 HGD-KVICESAIIVEYIDEVWSNNALS-----ILPQNAYDRANARFWVSYIDDKWLTSLK 113
                | : |||||:||:||::|    :  :     :|| : |:||  ||| ::|||:  :|
Db         61 LPDGRAICESAVIVQYIEDVARESGGAEAGSLLLPDDPYERAMHRFWTAFIDDKFWPALD 120

Qy        114 SVLATEDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYF-GGDTI-GFVDIGFGSFLSFI 171
              :|      |:    |     |  :|| |    | |:|:| |||    | :|:  | ||  :
Db        121 AVSLAPTPGARAQAAEDTRAALSLLEEAFKDRSNGRAFFSGGDAAPGLLDLALGCFLPAL 180

Qy        172 RVSENMNERKLLD--ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKW 225
              |  | ::   |:|   |  | |  |:: ||| || | :||:|||:|:|: : || ::
Db        181 RACERLHGLSLIDASATATPLLDGWSQRFAAHPAAKRVLPDTEKVVQFTRFLQAQF 236


RESULT 9
T10825
auxin-induced protein (clone MII-4) - mung bean (fragment)
C;Species: Vigna radiata (mung bean)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
C;Accession: T10825
R;Chen, J.; Wu, D.; Witham, F.H.; Heuser, C.W.; Arteca, R.N.
submitted to the EMBL Data Library, February 1995
A;Description: Molecular cloning and characterization of auxin-regulated genes from mung
A;Reference number: Z17176
A;Accession: T10825
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-230 <CHE>
A;Cross-references: EMBL:U20809; NID:g1184122; PID:g1184123
A;Experimental source: strain Rwilcz cv. Berken; tissue_type hypocotyl; clone MII-4
C;Superfamily: auxin-induced protein

  Query Match             34.4%;  Score 411;  DB 2;  Length 230;
  Best Local Similarity   44.4%;  Pred. No. 3.9e-26;
  Matches   92;  Conservative   31;  Mismatches   76;  Indels    8;  Gaps    4;

Qy         14 SPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVICESAIIVE 73
              |||  ||:||| |||::|: |||    ||| ||| |||||||:||| |||| : || :|||
Db         21 SPFVCRVKIALKLKGVEYKYVEENFRNKSEQLLKYNPVHKKVPVFVHGDKPLPESLVIVE 80

Qy         74 YIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWL-TSLKSVLATEDDEAKKLHFEQAE 132
              |||| |:||   ||  : | || ||||  :|||| :  | |||   :: | :|    .
Db         81 YIDETWNNN--PILASDPYQRALARFWSKFIDDKIVGASWKSVFTVDEKEREK----NIA 134

Qy        133 EVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLDETKYPGLT 192
              |  | :: : |:  | | :|||: :| |||       :| : : :   :||   |:| |
Db        135 ETYESLQFLENEIKE-KKFFGGEELGLVDIAAVYVAFWIPLIQEIAGLELLTSEKFPNLY 193

Qy        193 LWAETFAADPAVKGLLPETEKLVEFAK 219
                |:: |   | ||  ||  : :  | |
Db        194 RWSQEFLNHPIVKESLPPRDPVFAFFK 220


RESULT 10
S19185
parC protein - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C;Accession: S19185
R;Takahashi, Y.; Nagata, T.
submitted to the EMBL Data Library, February 1992
A;Description: Differential expression of an auxin-regulated gene, parC, and a novel rel
A;Reference number: S19182
A;Accession: S19185
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-221 <TAK>
A;Cross-references: EMBL:X64398; NID:g19905; PIDN:CAA45740.1; PID:g19906
C;Superfamily: auxin-induced protein

  Query Match             34.4%;  Score 410.5;  DB 2;  Length 221;
  Best Local Similarity   41.4%;  Pred. No. 4e-26;
  Matches   94;  Conservative   31;  Mismatches   91;  Indels   11;  Gaps    3;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||  :: ||  | | | :|::|||  | : ||  :| |  || |||: ||:||||||  |
Db          1 MANEEVILLDFWPSMFGMRLRIALAEKEIKYEYKQEDLRNKSPLLLQMNPIHKKIPVLIH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT-- 118
                | |||| | ||||:||| : | |:|| : |||| ||||  ||| |     : : ||
Db         61 NGKPICESIIAVEYIEEVWKDKAPSLLPSDPYDRAQARFWADYIDKKLYDFGRKLWATKG 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              |: || |  | :  :|||           : ||||:: |||||    | |:    |
Db        121 EEQEAAKKDFIECLKVLE-------GALGDRPYFGGESFGFVDIALIGFYSWFYAYETFG 173

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKW 225
                  | : |    ||:      :|   ||:  |::|| |:|: |:
Db        174 --NFSTEAECPKFVAWAKRCMQRESVAKSLPDQPKVLEFVKVLRQKF 218


RESULT 11
S16636
auxin-induced protein (clone pCNT107) - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Sep-1998
C;Accession: S16636
R;van der Zaal, E.J.; Droog, F.N.J.; Boot, C.J.M.; Hensgens, L.A.M.; Hoge, J.H.C.; Schil
Plant Mol. Biol. 16, 983-998, 1991
A;Title: Promoters of auxin-induced genes from tobacco can lead to auxin-inducible and r
A;Reference number: S16267; MUID:91322513; PMID:1863770
A;Accession: S16636
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-221 <ZAA>
A;Cross-references: EMBL:X56266
C;Superfamily: auxin-induced protein

  Query Match             34.1%;  Score 407.5;  DB 2;  Length 221;
  Best Local Similarity   40.3%;  Pred. No. 7.1e-26;
  Matches   94;  Conservative   29;  Mismatches   87;  Indels   23;  Gaps    4;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||  :: ||  | | | :|::|||  | : ||  || |  || |||: ||:||||||  |
Db          1 MANEEVILLDFWPSMFGMRLRIALAEKEIKYEYKEEDLRNKSPLLLQMNPIHKKIPVLIH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK--------WLTSL 112
                | |||| | ||||:||| : | ::|| : |||| ||||  ||| |         | |
Db         61 NGKPICESIIAVEYIEEVWKDKAPNLLPSDPYDRARARFWADYIDKKLYDFGRKLWTTK- 119
```

```
Qy        113 KSVLATEDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIR 172
Db        120 -----GEEQEAAKKDFIECLKVLE-------GALGDKPYFGGESFGFVDIALIGYYSWFY 167

Qy        173 VSENMNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKW 225
Db        168 AYETFG--NFSTEAECPKFVAWAKRCMQRESVAKSLPDQPKVLEFVKVLRQKF 218


RESULT 12
T09807
probable glutathione transferase (EC 2.5.1.18) - upland cotton (fragment)
N;Alternate names: probable glutathione S-transferase
C;Species: Gossypium hirsutum (upland cotton)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Nov-2000
C;Accession: T09807
R;Yamamoto, E.; Baird, W.V.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z16864
A;Accession: T09807
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-161 <YAM>
A;Cross-references: EMBL:AF064201; NID:g3135692; PID:g3135693
A;Experimental source: fiber
C;Genetics:
A;Gene: GST
C;Superfamily: auxin-induced protein
C;Keywords: transferase

  Query Match             33.7%;  Score 403;  DB 2;  Length 161;
  Best Local Similarity   51.0%;  Pred. No. 1.1e-25;
  Matches   78;  Conservative   28;  Mismatches   45;  Indels    2;  Gaps    1;

Qy         70 IIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAKKLHFE 129
Db          2 IIVQYIDEVWP--SAPILPSDPHERATARFWAAYLDDKWFPSLRAIGMAEGEDARKAAIG 59

Qy        130 QAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLDETKYP 189
Db         60 QVEEGLMLLEEAFGKCSQGQAFFGKDQIGYLDITFGCSLGWLRVTEKMSGIKLLNEINTP 119

Qy        190 GLTLWAETFAADPAVKGLLPETEKLVEFAKILQ 222
Db        120 ALLKWANRFCNDAAVKDVMPETEKLAEFAKMLR 152


RESULT 13
A36225
auxin-regulated protein, protoplast - common tobacco (cv. Xanthi nc)
N;Alternate names: gene str246C protein
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 17-Nov-2000
C;Accession: A36225; S52624; S52623; A36206; JQ0631; S47431; S47443
R;Takahashi, Y.; Kuroda, H.; Tanaka, T.; Machida, Y.; Takebe, I.; Nagata, T.
Proc. Natl. Acad. Sci. U.S.A. 86, 9279-9283, 1989
A;Title: Isolation of an auxin-regulated gene cDNA expressed during the transition from
A;Reference number: A36225; MUID:90083251; PMID:2594768
A;Accession: A36225
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-220 <TAK>
A;Cross-references: GB:M29274; NID:g170282; PIDN:AAA67894.1; PID:g170283
R;Froissard, D.; Gough, C.; Czernic, P.; Schneider, M.; Toppan, A.; Roby, D.; Marco, Y.
Plant Mol. Biol. 26, 515-521, 1994
A;Title: Structural organization of str 246C and str 246N, plant defense-related genes f
A;Reference number: S52623; MUID:95036027; PMID:7948901
A;Accession: S52624
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-220 <FRO>
A;Cross-references: EMBL:X80829; NID:g530739; PIDN:CAA56790.1; PID:g530740
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
A;Accession: S52623
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 48-194 <FRW>
A;Cross-references: EMBL:X80828; NID:g530737; PIDN:CAA56789.1; PID:g530738
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
R;Takahashi, Y.; Niwa, Y.; Machida, Y.; Nagata, T.
Proc. Natl. Acad. Sci. U.S.A. 87, 8013-8016, 1990
A;Title: Location of the cis-acting auxin-responsive region in the promoter of the par ge
A;Reference number: A36206; MUID:91045927; PMID:2236015
A;Accession: A36206
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-106 <TA2>
A;Cross-references: GB:D90215; NID:g218295; PIDN:BAA14243.1; PID:g218296
C;Genetics:
A;Gene: par
A;Introns: 106/3
C;Superfamily: auxin-induced protein

  Query Match             33.6%;  Score 402;  DB 2;  Length 220;
  Best Local Similarity   41.2%;  Pred. No. 2e-25;
  Matches   94;  Conservative   38;  Mismatches   78;  Indels   18;  Gaps    7;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
Db          1 MESNNVVLLDFWPSSFGMRLRIALALKGIKYEAKEENLSDKSPLLLEMNPVHKKIPILIH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLA--- 117
Db         61 NSKAICESLNILEYIDEVW-HDKCPLLPSDPYERSQARFWADYIDKKIYSTGRRVWSGKG 119

Qy        118 TEDDEAKKLHFEQAEEVLEKVE-EVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSEN 176
Db        120 EDQEEAKK----EFIEILKTLEGELGN-----KTYFGGDNLGFVDVALVPFTSWFYSYET 170

Qy        177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLK 224
Db        171 CANFSI--EAECPKLVVWAKTCMESESVSKSLPHPHKIYGF--VLELK 214


RESULT 14
S66354
glutathione transferase (EC 2.5.1.18), auxin-inducible - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 16-Feb-2001
C;Accession: S66354; S71497; E84696; S57711
R;van der Kop, D.A.M.; Schuyer, M.; Scheres, B.; van der Zaal, B.J.; Hooykaas, P.J.J.
Plant Mol. Biol. 30, 839-844, 1996
A;Title: Isolation and characterization of an auxin-inducible glutathione S-transferase
A;Reference number: S66354; MUID:96194465; PMID:8624414
A;Accession: S66354
A;Molecule type: DNA
A;Residues: 1-224 <VAN>
A;Cross-references: EMBL:X89216; NID:g895699; PIDN:CAA61504.1; PID:g895700
A;Accession: S71497
A;Molecule type: mRNA
A;Residues: 1-224 <VAX>
A;Cross-references: GB:X89216; NID:g895699; PIDN:CAA61504.1; PID:g895700
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-224 <STO>
A;Cross-references: GB:AE002093; NID:g3980390; PIDN:AAC95193.1; GSPDB:GN00139
C;Genetics:
A;Gene: At103-1a; At2g29450
```

```
A;Map position: 2
A;Introns: 106/3
C;Superfamily: auxin-induced protein
C;Keywords: transferase

  Query Match             33.4%;  Score 399;  DB 2;  Length 224;
  Best Local Similarity   41.1%;  Pred. No. 3.5e-25;
  Matches   90;  Conservative   39;  Mismatches   80;  Indels   10;  Gaps    5;

Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFF 59
              |||: :::||| | |||: ||::|| |||: || ||| |  || |||  ||:|||:||
Db          1 MAEKEEVKLLGIWASPFSRRVEMALKLKGIPYEYVEEILENKSPLLLALNPIHKKVPVLV 60

Qy         60 HGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATE 119
              |  | | || :|:||||| |  |   ||||: |:|: |||:   :|:: :      :|
Db         61 HNGKTILESHVILEYIDETWPQN--PILPQDPYERSKARFFAKLVDEQIMNVGFISMARA 118

Qy        120 DDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFI--RVSENM 177
              |:: :::  ||  |::  :|    |   || |||| |:||:|   || : |   |  | :
Db        119 DEKGREVLAEQVRELIMYLE----KELVGKDYFGGKTVGFLDFVAGSLIPFCLERGWEGI 174

Qy        178 NERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVE 216
                ::: | |:|    |         ||  :|  |: ||
Db        175 G-LEVITEEKFPEFKRWVRNLEKVEIVKDCVPPREEHVE 212


RESULT 15
G84696
probable glutathione S-transferase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84696
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-225 <STO>
A;Cross-references: GB:AE002093; NID:g3980388; PIDN:AAC95191.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g29470
A;Map position: 2
C;Superfamily: auxin-induced protein

  Query Match             33.3%;  Score 397.5;  DB 2;  Length 225;
  Best Local Similarity   38.9%;  Pred. No. 4.7e-25;
  Matches   88;  Conservative   54;  Mismatches   61;  Indels   23;  Gaps    9;

Qy          1 MAERD--LRLLGAWFSPFALRVQIALNLKGLDYEVVEET-LNPKSELLLKSNPVHKKIPV 57
              |||::  ::|:|:| |||: ||::|| |||: |: ::|  |  || |||: |||:||:||
Db          1 MAEKEEGVKLIGSWASPFSRRVEMALKLKGVPYDYLDEDYLVVKSPLLLQLNPVYKKVPV 60

Qy         58 FFHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK-WLTSLKSVL 116
                |  |:: || :|:|||: |:||   ||||: ||:| ||||  ::|::  :  |:|::
Db         61 LVHNGKILPESQLILEYIDQTWTNN--PILPQSPYDKAMARFWAKFVDEQVTMIGLRSLV 118

Qy        117 ATEDDEAKKLHFEQAEEVLEKVEEVF----NKCSEGKAYFGGDTIGFVDIGFGSFLSF-- 170
               :|          :: :  :|:|:|:     |: : ||  |||:||||:|:  || : |
Db        119 KSE---------KRIDVAIEEVQELIMLLENQIT-GKKLFGGETIGFLDMVVGSMIPFCL 168

Qy        171 IRVSENMNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVE 216
              |  | |    :: | |:| |  | :       |:  :|: || :|
Db        169 ARAWEGMG-IDMIPEEKFPELNRWIKNLKEIEIVRECIPDREKHIE 213


Search completed: October  8, 2004, 09:24:56
Job time : 40 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 8, 2004, 09:19:41 ; Search time 32 Seconds
(without alignments)
374.288 Million cell updates/sec

Title: US-10-088-945A-10
Perfect score: 1195
Sequence: 1 MAERDLRLLGAWFSPFALRV.....KLVEFAKILQLKWAAAAAAK 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1000.5	83.7	234	3	US-08-924-747-8	Sequence 8, Appli
2	1000.5	83.7	234	3	US-09-247-373B-8	Sequence 8, Appli
3	1000.5	83.7	234	3	US-09-296-715-8	Sequence 8, Appli
4	530.5	44.4	236	3	US-09-248-335-74	Sequence 74, Appl
5	512.5	42.9	232	3	US-09-248-335-56	Sequence 56, Appl
6	511	42.8	237	3	US-09-248-335-48	Sequence 48, Appl
7	504	42.2	225	3	US-09-248-335-62	Sequence 62, Appl
8	501	41.9	236	3	US-09-248-335-68	Sequence 68, Appl
9	489	40.9	228	3	US-09-248-335-70	Sequence 70, Appl
10	484.5	40.5	240	3	US-09-248-335-42	Sequence 42, Appl
11	481	40.3	231	3	US-09-248-335-54	Sequence 54, Appl
12	466	39.0	225	3	US-08-924-747-14	Sequence 14, Appl
13	466	39.0	225	3	US-09-247-373B-14	Sequence 14, Appl
14	466	39.0	225	3	US-09-296-715-14	Sequence 14, Appl
15	459.5	38.5	233	3	US-09-248-335-60	Sequence 60, Appl
16	458	38.3	225	3	US-09-248-335-40	Sequence 40, Appl
17	450.5	37.7	225	3	US-08-924-747-10	Sequence 10, Appl
18	450.5	37.7	225	3	US-09-247-373B-10	Sequence 10, Appl
19	450.5	37.7	225	3	US-09-296-715-10	Sequence 10, Appl
20	450	37.7	221	3	US-09-247-373B-54	Sequence 54, Appl
21	446	37.3	224	3	US-09-247-373B-34	Sequence 34, Appl
22	444	37.2	225	3	US-08-924-747-16	Sequence 16, Appl
23	444	37.2	225	3	US-09-247-373B-16	Sequence 16, Appl
24	444	37.2	225	3	US-09-296-715-16	Sequence 16, Appl
25	433.5	36.3	203	3	US-09-248-335-58	Sequence 58, Appl
26	428.5	35.9	229	3	US-09-247-373B-48	Sequence 48, Appl
27	428.5	35.9	235	3	US-09-248-335-64	Sequence 64, Appl
28	421.5	35.3	227	3	US-09-248-335-46	Sequence 46, Appl
29	410.5	34.4	220	3	US-08-924-747-26	Sequence 26, Appl
30	410.5	34.4	220	3	US-09-247-373B-26	Sequence 26, Appl
31	410.5	34.4	220	3	US-09-296-715-26	Sequence 26, Appl
32	409.5	34.3	219	3	US-09-247-373B-52	Sequence 52, Appl
33	409	34.2	222	3	US-09-248-335-66	Sequence 66, Appl
34	408	34.1	238	3	US-09-248-335-38	Sequence 38, Appl
35	404.5	33.8	222	3	US-08-924-747-22	Sequence 22, Appl
36	404.5	33.8	222	3	US-09-247-373B-22	Sequence 22, Appl
37	404.5	33.8	222	3	US-09-296-715-22	Sequence 22, Appl
38	404	33.8	200	3	US-09-248-335-72	Sequence 72, Appl
39	402	33.6	200	3	US-08-924-747-18	Sequence 18, Appl
40	402	33.6	200	3	US-09-247-373B-18	Sequence 18, Appl
41	402	33.6	200	3	US-09-296-715-18	Sequence 18, Appl
42	399.5	33.4	225	3	US-09-247-373B-36	Sequence 36, Appl
43	398	33.3	220	3	US-09-247-373B-50	Sequence 50, Appl
44	393	32.9	219	3	US-08-924-747-6	Sequence 6, Appli
45	393	32.9	219	3	US-08-924-747-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-924-747-8
; Sequence 8, Application US/08924747
; Patent No. 6063570
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; TITLE OF INVENTION: ENZYMES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,747
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SE3.03B09
US-08-924-747-8

Query Match 83.7%; Score 1000.5; DB 3; Length 234;
Best Local Similarity 84.6%; Pred. No. 1.8e-98;
Matches 192; Conservative 17; Mismatches 17; Indels 1; Gaps 1;

```
Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:||||||||||||:||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||||||||||||| |: | |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| |||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| |||||||||||||:||||:|||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 2
US-09-247-373B-8
; Sequence 8, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
;  APPLICANT:  MCGONIGLE, BRIAN
;  APPLICANT:  O'KEEFE, DANIEL
;  TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE:  CL-1108-A
;  CURRENT APPLICATION NUMBER: US/09/247,373B
;  CURRENT FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  08/924,747
;  PRIOR FILING DATE:  1997-09-05
;  NUMBER OF SEQ ID NOS:  56
;  SOFTWARE:  Microsoft Office 97
; SEQ ID NO 8
;   LENGTH: 234
;   TYPE: PRT
;   ORGANISM: SOYBEAN
US-09-247-373B-8

  Query Match             83.7%;  Score 1000.5;  DB 3;  Length 234;
  Best Local Similarity   84.6%;  Pred. No. 1.8e-98;
  Matches  192;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:||||||||||||:||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||||||||||||| |: | |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| |||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| |||||||||||||:||||:|||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 3
US-09-296-715-8
; Sequence 8, Application US/09296715
; Patent No. 6171839
;  GENERAL INFORMATION:
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/296,715
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  234 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  SOYBEAN
;    IMMEDIATE SOURCE:
;      CLONE:  SE3.03B09
US-09-296-715-8

  Query Match             83.7%;  Score 1000.5;  DB 3;  Length 234;
  Best Local Similarity   84.6%;  Pred. No. 1.8e-98;
  Matches  192;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:||||||||||||:||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||||||||||||| |: | |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| |||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| |||||||||||||:||||:|||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 4
US-09-248-335-74
; Sequence 74, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 74
;   LENGTH: 236
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-74
```

```
  Query Match             44.4%;  Score 530.5;  DB 3;  Length 236;
  Best Local Similarity   47.2%;  Pred. No. 2.1e-48;
  Matches  109;  Conservative   48;  Mismatches   65;  Indels    9;  Gaps    4;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLN-PKSELLLKSNPVHKKIPVFFHGDK 63
              :|:||||| ||| |||::||: ||| || ||| |:  ||||||:|||||||:|| | |
Db          8 ELKLLGAWASPFVLRVKLALSFKGLSYEDVEEDLSGGKSELLLESNPVHKKVPVLLHNGK 67

Qy         64 VICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK----WLTSLKSVLATE 119
               :||| |||:|||| ::    |:|| : : || |||| :|||||    || | ::   |:
Db         68 PVCESQIIVQYIDEAFAGTGPSLLPADPHQRAVARFWGAYIDDKLLAFWLQSARA--KTQ 125

Qy        120 DDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNE 179
              :::|: |  :||    | :|  | : |||| :||||::|::|:  |: ::::  :| :
Db        126 EEKAEAL--KQALAAAENLEAAFTEISEGKPFFGGDSVGYLDVTLGALVAWVHAAEKLYG 183

Qy        180 RKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              :| | |: | |: : | | |  | | :||: : |||:||  |   ||||:
Db        184 MRLFDATRTPRLSAFVERFGALGAAKAVLPDVDGLVEYAKQRQADAAAAAS 234


RESULT 5
US-09-248-335-56
; Sequence 56, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 56
;   LENGTH: 232
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-56

  Query Match             42.9%;  Score 512.5;  DB 3;  Length 232;
  Best Local Similarity   46.3%;  Pred. No. 1.7e-46;
  Matches  101;  Conservative   41;  Mismatches   75;  Indels    1;  Gaps    1;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |:|  :|::| | ||| :|| ||| |||:::| ||| :  |||||||||||||||  |
Db          1 MSEAAVRVIGLWPSPFVIRVLIALKLKGVEFEFVEEVVGRKSELLLRSNPVHKKIPVLLH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               | | || |:|:|||||||: | : || :|: ||  |||  |:|||   ::: :|   |
Db         61 HGKPISESLIVVQYIDEVWSSGAPAFLPVDAHARAVQRFWAQYVDDKLPWAIR-ILKGTD 119

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              |   :    |    |: :|| | : |:|| |||||::|::||   | : :::  | :
Db        120 DGGMEQAAGQLSAALQLLEEAFAQLSQGKRYFGGDSVGYLDIALVSHVGWVKAVEKIAGV 179

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFA 218
              |||: .| | |  ||:   | |||   :|: :| |||:
Db        180 TLLDKAKVPNLVAWADRLCAHPAVVDAIPDADKFVEFS 217


RESULT 6
US-09-248-335-48
; Sequence 48, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 48
;   LENGTH: 237
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-48

  Query Match             42.8%;  Score 511;  DB 3;  Length 237;
  Best Local Similarity   44.7%;  Pred. No. 2.5e-46;
  Matches  101;  Conservative   41;  Mismatches   84;  Indels    0;  Gaps    0;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              |:||||| |||  ||::||:||||:|| ||| |  ||:||| |||||| :||  |||| |
Db         11 LKLLGAWPSPFVNRVRMALHLKGLEYENVEEDLTNKSDLLLASNPVHKLVPVLLHGDKPI 70

Qy         66 CESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAKK 125
               || :||||:|: :     ::|| : |:|| ||||  |:| |    :  |  :| :
Db         71 SESLVIVEYLDDAFPGAGQAVLPADPYERAVARFWAKYVDGKLHGMMVKALMGATEEERA 130

Qy        126 LHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLDE 185
                  |   :: :|  | :|| || :| ||  |::|:  | |: ::|  : :   ||||
Db        131 TATVDALAAMDTLEGAFAECSGGKKFFAGDAPGYLDVALGGFIGWLRAWDKVGGVKLLDA 190

Qy        186 TKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAA 231
              : | |  ||| |||    | ::|: : : ||||:|| : |||| :
Db        191 GRVPRLATWAERFAALDVAKEVIPDPDHIAEFAKVLQARSAAAATS 236


RESULT 7
US-09-248-335-62
; Sequence 62, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 62
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-62

  Query Match             42.2%;  Score 504;  DB 3;  Length 225;
  Best Local Similarity   46.6%;  Pred. No. 1.3e-45;
  Matches  102;  Conservative   38;  Mismatches   77;  Indels    2;  Gaps    2;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              |:|  :|::| | ||| :|| ||| || ::|| ||| :  |||||| ||||||||||  |
Db          1 MSEAAVRVIGLWPSPFVIRVLIALKLKHVEYEFVEEVVGSKSELLLASNPVHKKIPVLLH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNA-LSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATE 119
               | : || |||:|||||||: |  :||| : | ||  |||  |:||   ::: :  |
Db         61 HGKPLSESLIIVQYIDEVWSSGAPAAILPADPYARAVQRFWAQYVDDKMHPAIRVLKGTY 120

Qy        120 DDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNE 179
              | : |:    |    |: :|| | :  :|| |||||::|::||   | : :::  | :
Db        121 DGD-KEQAAGQLSAALQLLEEAFAQLGQGKRYFGGDSVGYLDIALVSHVGWVKAVEKIAG 179

Qy        180 RKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFA 218
              |||| | | |  ||:   | |||   :|: :| |||:
Db        180 VTLLDEAKVPNLVAWADRLCAHPAVVDAIPDADKFVEFS 218
```

```
RESULT 8
US-09-248-335-68
; Sequence 68, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 68
;   LENGTH: 236
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-68

  Query Match             41.9%;  Score 501;  DB 3;  Length 236;
  Best Local Similarity   45.5%;  Pred. No. 2.8e-45;
  Matches  102;  Conservative   39;  Mismatches   79;  Indels    4;  Gaps    2;

Qy          1 MAERD--LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVF 58
              :|::|  |::|| | ||| :| ::|||||||  |  ||: |: |||||| ||||| |:||
Db          5 VADKDPELKVLGVWSSPFVIRARVALNLKGLAYRYVEDNLDSKSELLLASNPVHGKVPVL 64

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
               |  : :||| :|||||| :  :   :|| : | ||  ||| || |||   :   |
Db         65 LHDGRPVCESRVIVEYIDEAFPASGPCLLPADPYRRAVDRFWASYADDKLFPTWIPVYNG 124

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKC--SEGKAYFGGDTIGFVDIGFGSFLSFIRVSEN 176
                 | :     |   |||| |: |::|  | |||:|||  | ||:  | || ::| ||
Db        125 RTSEDRVAAARQVVAVLEKFEQAFDECSGSGGKAFFGGDAAGLVDVVLGGFLGWLRASEA 184

Qy        177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKI 220
              |   :::|  | | |  ||: |||   |: ::|: ::|:|: ||
Db        185 MCGVRVIDPAKTPLLAAWADRFAALDGVREVIPDVQRLLEYNKI 228


RESULT 9
US-09-248-335-70
; Sequence 70, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 70
;   LENGTH: 228
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-70

  Query Match             40.9%;  Score 489;  DB 3;  Length 228;
  Best Local Similarity   47.4%;  Pred. No. 5.1e-44;
  Matches  102;  Conservative   31;  Mismatches   80;  Indels    2;  Gaps    2;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              :|:||| | ||:  |||| |||||| || ||| |  |||||| |||||||:||  |  |
Db          8 ELKLLGVWDSPYVNRVQIVLNLKGLSYEYVEEDLLSKSELLLNSNPVHKKVPVLIHAGKP 67

Qy         65 ICESAIIVEYIDEVWSNNAL-SILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEA 123
              : ||  ||:|:|| : :    |:||   | || |||| :::|||  :   :||   :
Db         68 VAESQAIVQYLDEAFPSGTFPSVLPAEPYARATARFWAAFVDDKVGSPWHTVLFAREHGK 127

Qy        124 KKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLL 183
              |      :    || :|  |  || |: ||||| |||||:  ||:| : :| | |   ::|
Db        128 KADAASRIVAALETLEGAFEDCSGGRDYFGGDAIGFVDVVLGSYLGWFKVFEKMVGVRVL 187

Qy        184 DETKYPGLTLWAETFAADPAVKGLLP-ETEKLVEF 217
              |  : | |  | | |||  | | :|| : :|::||
Db        188 DVARTPLLAAWGERFAAAEAAKDVLPDDVDKVLEF 222


RESULT 10
US-09-248-335-42
; Sequence 42, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 42
;   LENGTH: 240
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-42

  Query Match             40.5%;  Score 484.5;  DB 3;  Length 240;
  Best Local Similarity   44.0%;  Pred. No. 1.7e-43;
  Matches  102;  Conservative   39;  Mismatches   78;  Indels   13;  Gaps    4;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              | |||   |||||||::||:|||| || :|: |  | |||| |||||||:||  |  | |
Db         11 LTLLGLHVSPFALRVRMALSLKGLSYEYIEQDLFHKGELLLSSNPVHKKVPVLIHHGKPI 70

Qy         66 CESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK----WLTSLKSVLATEDD 121
              |||  :|||:||||   | :||| : : || |||| :||| |    |  :|:  |||:
Db         71 CESLAVVEYVDEVWPGAAATILPADPHGRATARFWAAYIDGKLFPAWTGIMKA--ATEEA 128

Qy        122 EAKKLHFEQAEEVLEKVEEVFNKCS-----EGKAYFGGDTIGFVDIGFGSFLSFIRVSEN 176
                | ||   :    :  :|: | : |      :| |:||||::|::|:  |  | :
Db        129 RADKL--RETHAAVLNLEKAFAEISSSSSNDGAAFFGGDSVGYLDLALGCSLPWFGALRA 186

Qy        177 MNERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAA 228
              |   :::|  : | |  ||| |   |  | :||: :: | :|| :|  ||:|
Db        187 MLGVEIIDAAQAPLLVAWAERFGETPVAKEVLPQPDEAVAYAKKIQAYWASA 238


RESULT 11
US-09-248-335-54
; Sequence 54, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 54
;   LENGTH: 231
;   TYPE: PRT
```

```
;   ORGANISM: maize
US-09-248-335-54

  Query Match             40.3%;  Score 481;  DB 3;  Length 231;
  Best Local Similarity   43.5%;  Pred. No. 3.7e-43;
  Matches   97;  Conservative   38;  Mismatches   80;  Indels    8;  Gaps    3;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||::|| | ||| :||:| |||||| || ||| |  || ||| |||||| :||  |  :
Db          6 DLKVLGVWTSPFVIRVRIVLNLKGLAYEYVEEDLGNKSALLLGSNPVHKSVPVLLHAGRA 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDK----WLTSLKSVLATED 120
              | || :|::|||||:    ::|| : |:|| |||| :|||||    ||  |       :
Db         66 INESQVILQYIDEVWAGTGPAVLPADPYERAVARFWGAYIDDKVESAWLGML---FRCAN 122

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              :| :     :| | |: :|  |  || |: :|||| |||||    | :| :      :
Db        123 EEERAAAVARAREALDALEGAFRDCSRGRPFFGGDDIGFVDAVLGGYLGWFGAVGRIIGS 182

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLP-ETEKLVEFAKILQ 222
              :|:|  : | |  | : | |    ||::| : :|:: | : |:
Db        183 RLIDPARTPLLAAWEDRFRAADVAKGVVPDDLDKMLAFLQTLR 225


RESULT 12
US-08-924-747-14
; Sequence 14, Application US/08924747
; Patent No. 6063570
;  GENERAL INFORMATION:
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/924,747
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  SOYBEAN
;    IMMEDIATE SOURCE:
;      CLONE:  SES8W.PK0028.C6
US-08-924-747-14

  Query Match             39.0%;  Score 466;  DB 3;  Length 225;
  Best Local Similarity   45.9%;  Pred. No. 1.4e-41;
  Matches  100;  Conservative   37;  Mismatches   73;  Indels    8;  Gaps    4;

Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGD 62
              : |::|||  |||  |||||| |||: |: :|| |  |||||||||||||:||| | :
Db          5 QEDVKLLGATGSPFVCRVQIALKLKGVQYKFLEENLRNKSELLLKSNPVHKKVPVFIHNE 64

Qy         63 KVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDE 122
              | | || :|||||| | ||   ||| : | || |||  :|||| : :    : | |::
Db         65 KPIAESLVIVEYIDETWKNN--PILPSDPYQRALARFWSKFIDDKVVGAAWKYIYTVDEK 122

Qy        123 AKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSF-IRVSENMNERK 181
              :: : |:: | |: :|       : | :|||: || |||    |::| | : : :   |
Db        123 EREKNVEESYEALQFLENEL----KDKKFFGGEEIGLVDIA-AVFIAFWIPIIQEVLGLK 177

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              |    |:| |  |:: |   | || :||  ::|  | |
Db        178 LFTSEKFPKLYKWSQEFINHPVVKQVLPPRDQLFAFYK 215


RESULT 13
US-09-247-373B-14
; Sequence 14, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
;  APPLICANT:  MCGONIGLE, BRIAN
;  APPLICANT:  O'KEEFE, DANIEL
;  TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE:  CL-1108-A
;  CURRENT APPLICATION NUMBER: US/09/247,373B
;  CURRENT FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  08/924,747
;  PRIOR FILING DATE:  1997-09-05
;  NUMBER OF SEQ ID NOS:  56
;  SOFTWARE:  Microsoft Office 97
; SEQ ID NO 14
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: SOYBEAN
US-09-247-373B-14

  Query Match             39.0%;  Score 466;  DB 3;  Length 225;
  Best Local Similarity   45.9%;  Pred. No. 1.4e-41;
  Matches  100;  Conservative   37;  Mismatches   73;  Indels    8;  Gaps    4;

Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGD 62
              : |::|||  |||  |||||| |||: |: :|| |  |||||||||||||:||| | :
Db          5 QEDVKLLGATGSPFVCRVQIALKLKGVQYKFLEENLRNKSELLLKSNPVHKKVPVFIHNE 64

Qy         63 KVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDE 122
              | | || :|||||| | ||   ||| : | || |||  :|||| : :    : | |::
Db         65 KPIAESLVIVEYIDETWKNN--PILPSDPYQRALARFWSKFIDDKVVGAAWKYIYTVDEK 122

Qy        123 AKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSF-IRVSENMNERK 181
              :: : |:: | |: :|       : | :|||: || |||    |::| | : : :   |
Db        123 EREKNVEESYEALQFLENEL----KDKKFFGGEEIGLVDIA-AVFIAFWIPIIQEVLGLK 177

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              |    |:| |  |:: |   | || :||  ::|  | |
Db        178 LFTSEKFPKLYKWSQEFINHPVVKQVLPPRDQLFAFYK 215


RESULT 14
US-09-296-715-14
; Sequence 14, Application US/09296715
; Patent No. 6171839
;  GENERAL INFORMATION:
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
```

```
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/296,715
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  SOYBEAN
;    IMMEDIATE SOURCE:
;      CLONE:  SES8W.PK0028.C6
US-09-296-715-14

  Query Match             39.0%;  Score 466;  DB 3;  Length 225;
  Best Local Similarity   45.9%;  Pred. No. 1.4e-41;
  Matches  100;  Conservative   37;  Mismatches   73;  Indels    8;  Gaps    4;

Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGD 62
              : |::||||  |||  |||||| |||: |: :|| |  |||||||||||||:||| | :
Db          5 QEDVKLLGATGSPFVCRVQIALKLKGVQYKFLEENLRNKSELLLKSNPVHKKVPVFIHNE 64

Qy         63 KVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDE 122
               | | || :|||||| | ||   ||| : | || ||||  :|||| : :    : | |::
Db         65 KPIAESLVIVEYIDETWKNN--PILPSDPYQRALARFWSKFIDDKVVGAAWKYIYTVDEK 122

Qy        123 AKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSF-IRVSENMNERK 181
               :: : |:: | |: :|       : | :|||: || |||    |::| | : : :   |
Db        123 EREKNVEESYEALQFLENEL----KDKKFFGGEEIGLVDIA-AVFIAFWIPIIQEVLGLK 177

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              |    |:| |  |:: |   | || :||  ::|  | |
Db        178 LFTSEKFPKLYKWSQEFINHPVVKQVLPPRDQLFAFYK 215


RESULT 15
US-09-248-335-60
; Sequence 60, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 60
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: maize
US-09-248-335-60

  Query Match             38.5%;  Score 459.5;  DB 3;  Length 233;
  Best Local Similarity   43.6%;  Pred. No. 7.3e-41;
  Matches   99;  Conservative   40;  Mismatches   83;  Indels    5;  Gaps    3;

Qy          6 LRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKVI 65
              |::||   ||| ||| :||| ||: || ||| |: |||||||||||||:||  |  | |
Db          7 LKVLGLQVSPFVLRVCMALNTKGVSYEYVEEDLSNKSELLLKSNPVHKKVPVLIHNGKPI 66

Qy         66 CESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAKK 125
              ||| :|::|:||:::    |||| : |:|| |||| :| ||| | :   ::  :  | :
Db         67 CESLVIMQYVDELFAGR--SILPTDPYERATARFWAAYADDKLLPAWYGMVKAQSAEERA 124

Qy        126 LHFEQAEEVLEKVEEVFNKCSEGK-AYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
                |:    :: :|  | ||| |  |:||||:||:|||  |||| :      : : ::::
Db        125 EKVEETLSAIQHMEVAFAKCSGGNAAFFGGDSIGYVDIVLGSFLFWFEAVRRVYDLEIIN 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              :  | |  ||| |      | ::|  :  |: |     |  | |||
Db        185 ASNTPLLAAWAERFVGTVEAKEVVPVPD--VDMAVQCINKLHAPAAA 229


Search completed: October  8, 2004, 09:25:33
Job time : 33 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 8, 2004, 09:04:30 ; Search time 123 Seconds
(without alignments)
532.935 Million cell updates/sec

Title: US-10-088-945A-10
Perfect score: 1195
Sequence: 1 MAERDLRLLGAWFSPFALRV.....KLVEFAKILQLKWAAAAAAK 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1195	100.0	232	4	AAB62226	Aab62226 Glycine m
2	1000.5	83.7	234	3	AAB03734	Aab03734 Clone se3
3	1000.5	83.7	234	3	AAY79515	Aay79515 Soybean g
4	1000.5	83.7	234	3	AAB07829	Aab07829 Amino aci
5	1000.5	83.7	234	4	AAB66733	Aab66733 Soybean t
6	1000.5	83.7	234	4	AAB62227	Aab62227 Glycine m
7	638	53.4	227	5	ABB90937	Abb90937 Herbicida
8	624	52.2	233	4	AAB47340	Aab47340 GST3. 9/2
9	623	52.1	233	4	AAB47339	Aab47339 GST2. 9/2
10	607	50.8	233	4	AAB47341	Aab47341 GST conse
11	596	49.9	234	5	ABB91390	Abb91390 Herbicida
12	585.5	49.0	234	3	AAG26676	Aag26676 Arabidops
13	585.5	49.0	281	3	AAG26675	Aag26675 Arabidops
14	585.5	49.0	287	3	AAG26674	Aag26674 Arabidops
15	558	46.7	233	3	AAG53029	Aag53029 Arabidops
16	558	46.7	233	5	ABB91389	Abb91389 Herbicida
17	558	46.7	241	3	AAG53028	Aag53028 Arabidops
18	554	46.4	233	3	AAG11459	Aag11459 Arabidops
19	554	46.4	241	3	AAG11458	Aag11458 Arabidops
20	530.5	44.4	236	3	AAB22136	Aab22136 Maize glu
21	526	44.0	227	3	AAG08276	Aag08276 Arabidops
22	526	44.0	245	3	AAG08275	Aag08275 Arabidops
23	523.5	43.8	233	2	AAY05544	Aay05544 Wheat Typ
24	520	43.5	227	3	AAG43149	Aag43149 Arabidops
25	520	43.5	227	5	ABB91146	Abb91146 Herbicida
26	512.5	42.9	232	3	AAB22127	Aab22127 Maize glu
27	511	42.8	237	3	AAB22123	Aab22123 Maize glu
28	504	42.2	225	3	AAB22130	Aab22130 Maize glu
29	501	41.9	236	3	AAB22133	Aab22133 Maize glu
30	491.5	41.1	234	5	ABB91525	Abb91525 Herbicida
31	489	40.9	228	3	AAB22134	Aab22134 Maize glu
32	486.5	40.7	229	5	ABB91524	Abb91524 Herbicida
33	484.5	40.5	240	3	AAB22120	Aab22120 Maize glu
34	481	40.3	231	3	AAB22126	Aab22126 Maize glu
35	479.5	40.1	243	5	ABB91147	Abb91147 Herbicida
36	466	39.0	225	3	AAB03737	Aab03737 Clone ses
37	466	39.0	225	3	AAY79518	Aay79518 Soybean g
38	466	39.0	225	3	AAB07832	Aab07832 Amino aci
39	466	39.0	225	4	AAB66736	Aab66736 Soybean t
40	465	38.9	170	5	ABB90938	Abb90938 Herbicida
41	459.5	38.5	233	3	AAB22129	Aab22129 Maize glu
42	458	38.3	225	3	AAB22119	Aab22119 Maize glu
43	450.5	37.7	225	3	AAB03735	Aab03735 Clone se6
44	450.5	37.7	225	3	AAY79516	Aay79516 Soybean g
45	450.5	37.7	225	3	AAB07830	Aab07830 Amino aci

ALIGNMENTS

RESULT 1
AAB62226
ID AAB62226 standard; protein; 232 AA.
XX
AC AAB62226;
XX
DT 11-JUN-2001 (first entry)
XX
DE Glycine max glutathione-S-transferase (GST) 3.3.
XX
KW Soybean; glutathione-S-transferase; GST; homoglutathione synthetase; HGS;
KW beta-alanine; gamma-glutamylcysteine; herbicide.
XX
OS Glycine max.
XX
PN WO200121770-A2.
XX
PD 29-MAR-2001.
XX
PF 18-SEP-2000; 2000WO-GB003573.
XX
PR 21-SEP-1999; 99GB-00022346.
XX
PA (ZENE) ZENECA LTD.
XX
PI Andrews CJ, Jepson I, Townson JK, Edwards R, Cummins I;
PI Skipsey M;
XX
DR WPI; 2001-257978/26.
DR N-PSDB; AAF57513.
XX
PT Novel glutathione-S-transferase and homoglutathione synthetase sequences
PT from soybean for producing plants which are resistant and tolerant to
PT herbicide comprising fomesafen and/or acifluorfen.
XX
PS Claim 1; Page 44-45; 64pp; English.
XX
CC The invention relates to new soybean glutathione-S-transferase (GST) and
CC a homoglutathione synthetase (HGS). The HGS is capable of catalyzing the
CC addition of beta-alanine onto gamma-glutamylcysteine. Polynucleotides
CC encoding the proteins of the invention are useful for producing plants
CC which are resistant and/or tolerant to a herbicide comprising fomesafen
CC and/or acifluorfen. Methods of the invention are useful for providing
CC plants with further desired agronomic trait, especially resistant to a
CC herbicide, comprising glyphosate or its salt. Further desired agronomic
CC traits include insect resistance, nematode resistance, stress tolerance,
CC altered field, altered nutritional value, altered quality or any other
CC desirable agronomic trait. GST or its variant is also useful as a

```
CC   selectable marker gene. The present sequence represents a Glycine max GST
CC   3.3 protein (also referred to as GST 3.6)
XX
SQ   Sequence 232 AA;

  Query Match           100.0%;  Score 1195;  DB 4;  Length 232;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-118;
  Matches  232;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAK 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAAK 232


RESULT 2
AAB03734
ID   AAB03734 standard; protein; 234 AA.
XX
AC   AAB03734;
XX
DT   04-OCT-2000  (first entry)
XX
DE   Clone se3.03b09 type III GST protein sequence.
XX
KW   Soybean; glutathione-S-transferase; GST; detoxify; herbicide; stress;
KW   transgenic plant; tolerant; plant breeding.
XX
OS   Glycine max.
XX
PN   US6063570-A.
XX
PD   16-MAY-2000.
XX
PF   05-SEP-1997;   97US-00924747.
XX
PR   05-SEP-1997;   97US-00924747.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Mcgonigle B,  O'keefe DP;
XX
DR   WPI; 2000-375487/32.
DR   N-PSDB; AAA53398.
XX
PT   New Glutathione-S-Transferase enzymes and isolated nucleic acid fragments
PT   encoding them, useful for detoxifying xenobiotic compounds in plants and
PT   seeds, as well as in producing transgenic plants that are herbicide-
PT   resistant.
XX
PS   Claim 1; Col 35-36; 36pp; English.
XX
CC   This sequence represents a Glutathione-S-Transferase (GST) protein
CC   isolated from a soybean clone. The invention relates to isolated nucleic
CC   acid fragments (see AAA53393-A53406) which encode soybean GST
CC   polypeptides (AAB03731-B03744). GSTs are a family of enzymes which
CC   catalyse the conjugation of glutathione, homoglutathione and other
CC   glutathione-like analogues, to a large range of hydrophobic,
CC   electrophilic compounds. GSTs have been implicated in the detoxification
CC   of certain herbicides. The GST nucleotide sequences are useful in the
CC   construction of herbicide-tolerant transgenic plants, plants that are
CC   tolerant to a wide variety of stresses, or plants in which the GST
CC   enzymes are present at higher or lower levels than they are normally. The
CC   nucleic acid fragments are also useful as probes for genetically and
CC   physically mapping the genes that they are part of, and as markers for
CC   traits linked to expression of the enzymes. This will be useful in plant
CC   breeding in order to develop lines with desired phenotypes or in the
CC   identification of mutants. The soybean GST enzymes are used to detoxify
CC   xenobiotic compounds in plants and seeds. The enzymes are also useful as
CC   targets to facilitate design and/or identify inhibitors of the enzymes
CC   that may be used as herbicides or herbicide synergists. The GST enzymes
CC   produced in the host cells, particularly in microbial host cells, are
CC   useful in preparing antibodies to the enzymes. These antibodies are
CC   useful for detecting the enzymes in situ in cells or in vitro in cell
CC   extracts
XX
SQ   Sequence 234 AA;

  Query Match           83.7%;  Score 1000.5;  DB 3;  Length 234;
  Best Local Similarity  84.6%;  Pred. No. 1.1e-97;
  Matches  192;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:||||||||||:||||||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||||||||||| |: | |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| ||||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| ||||||||||||:||||:|||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 3
AAY79515
ID   AAY79515 standard; protein; 234 AA.
XX
AC   AAY79515;
XX
DT   01-AUG-2000  (first entry)
XX
DE   Soybean glutathione-S-transferase se3.03B09.
XX
KW   Soybean; glutathione-S-transferase; GST; xenobiotic; detoxification;
KW   transgenic plant; herbicide tolerance.
XX
OS   Glycine max.
XX
PN   WO200018936-A1.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1998;   98WO-US020501.
XX
PR   30-SEP-1998;   98WO-US020501.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Mcgonigle B,  O'keefe DP;
XX
DR   WPI; 2000-317517/27.
DR   N-PSDB; AAZ94952.
XX
PT   Nucleic acids encoding soybean glutathione-S-transferase enzymes useful
PT   for conferring herbicide resistance to plants.
XX
PS   Claim 4; Page 46; 76pp; English.
XX
CC   The present sequence is that a soybean class III glutathione-S-
```

CC transferase (GST), as deduced from soybean embryo cDNA clone se3.03B09
CC (see AAZ94952). The invention provides soybean GST enzymes (see AAY79512-
CC 25) involved in the detoxification of xenobiotic compounds, especially
CC herbicides, in plants and seeds. Chimeric genes encoding all or a portion
CC of soybean GST, host cells, and methods of recombinant production of
CC soybean GST enzymes are provided. The sequences are useful in the
CC construction of herbicide-tolerant transgenic plants, in the recombinant
CC production of GST enzymes, in the development of screening assays to
CC identify compounds inhibitory to the GST enzymes (useful as herbicides or
CC herbicide synergists), and in screening assays to identify chemical
CC substrates of the GSTs
XX
SQ Sequence 234 AA;

Query Match 83.7%; Score 1000.5; DB 3; Length 234;
Best Local Similarity 84.6%; Pred. No. 1.1e-97;
Matches 192; Conservative 17; Mismatches 17; Indels 1; Gaps 1;

```
Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:|||||||||||:|||||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              ||||||||||||| |: |  |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| ||||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| ||||||||||||:||||:|||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231
```

RESULT 4
AAB07829
ID AAB07829 standard; protein; 234 AA.
XX
AC AAB07829;
XX
DT 14-NOV-2000 (first entry)
XX
DE Amino acid sequence of a soybean type III glutathione-S-transferase.
XX
KW Soybean; glutathione-S-transferase; GST; detoxification;
KW xenobiotic compound; herbicide-tolerance; transgenic plant;
KW herbicide synergist.
XX
OS Glycine max.
XX
PN WO200047728-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US003347.
XX
PR 10-FEB-1999; 99US-00247373.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Mcgonigle B, O'keefe DP;
XX
DR WPI; 2000-549144/50.
DR N-PSDB; AAA59460.
XX
PT Soybean glutathione-S-transferase polypeptides and polynucleotides used
PT to produce herbicide tolerant transgenic plants and to screen for
PT inhibitors or substrates of the enzyme.
XX
PS Claim 4; Page 58-59; 84pp; English.
XX
CC The present sequence represents a soybean glutathione-S-transferase (GST)
CC enzyme. The enzyme is involved in the detoxification of xenobiotic
CC compounds in plants and seeds. The GST polynucleotides and polypeptides
CC are used for the production of herbicide-tolerant transgenic plants, and
CC for the development of screening assays to identify GST inhibitors and
CC substrates, which can be used as herbicide synergists. GST Gene specific
CC probes can be used in gene identification methods. The recombinant GST
CC enzymes can be used to produce enzyme specific antibodies which are used
CC to detect the enzymes in situ in cells or in vitro in cell extracts
XX
SQ Sequence 234 AA;

Query Match 83.7%; Score 1000.5; DB 3; Length 234;
Best Local Similarity 84.6%; Pred. No. 1.1e-97;
Matches 192; Conservative 17; Mismatches 17; Indels 1; Gaps 1;

```
Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:|||||||||||:|||||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              ||||||||||||| |: |  |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| ||||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| ||||||||||||:||||:|||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231
```

RESULT 5
AAB66733
ID AAB66733 standard; protein; 234 AA.
XX
AC AAB66733;
XX
DT 09-APR-2001 (first entry)
XX
DE Soybean type III GST protein #2.
XX
KW Soybean; glutathione-S-transferase; herbicide; GST.
XX
OS Glycine max.
XX
PN US6171839-B1.
XX
PD 09-JAN-2001.
XX
PF 22-APR-1999; 99US-00296715.
XX
PR 05-SEP-1997; 97US-00924747.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Mcgonigle B, O'keefe DP;
XX
DR WPI; 2001-136874/14.
XX
PT Novel soybean glutathione-S-transferase enzymes useful as targets to
PT facilitate design and/or identification of inhibitors of the enzyme, that
PT are used as herbicides or herbicide synergists.
XX
PS Claim 1; Col 35-38; 37pp; English.
XX
CC The present invention relates to soybean glutathione-S-transferase
CC proteins. The novel sequences are useful in the construction of herbicide
CC tolerant transgenic plants, in the recombinant production of glutathione-
CC S-transferase (GST) enzymes, in the development of screening assays to
CC identify compounds inhibitory to the GST enzymes, and in screening assays
CC to identify chemical substrates of the GSTs
XX

```
SQ   Sequence 234 AA;

  Query Match             83.7%;  Score 1000.5;  DB 4;  Length 234;
  Best Local Similarity   84.6%;  Pred. No. 1.1e-97;
  Matches  192;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:||||||||||:||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||||||||||| |: | |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| ||||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| |||||||||||||:||||:|||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 6
AAB62227
ID   AAB62227 standard; protein; 234 AA.
XX
AC   AAB62227;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Glycine max clone SE3.03B09 sequence.
XX
KW   Soybean; glutathione-S-transferase; GST; homoglutathione synthetase; HGS;
KW   beta-alanine; gamma-glutamylcysteine; herbicide.
XX
OS   Glycine max.
XX
PN   WO200121770-A2.
XX
PD   29-MAR-2001.
XX
PF   18-SEP-2000; 2000WO-GB003573.
XX
PR   21-SEP-1999;   99GB-00022346.
XX
PA   (ZENE ) ZENECA LTD.
XX
PI   Andrews CJ,  Jepson I,  Townson JK,  Edwards R,  Cummins I;
PI   Skipsey M;
XX
DR   WPI; 2001-257978/26.
DR   N-PSDB; AAF57526.
XX
PT   Novel glutathione-S-transferase and homoglutathione synthetase sequences
PT   from soybean for producing plants which are resistant and tolerant to
PT   herbicide comprising fomesafen and/or acifluorfen.
XX
PS   Disclosure; Page 59-60; 64pp; English.
XX
CC   The invention relates to new soybean glutathione-S-transferase (GST) and
CC   a homoglutathione synthetase (HGS). The HGS is capable of catalyzing the
CC   addition of beta-alanine onto gamma-glutamylcysteine. Polynucleotides
CC   encoding the proteins of the invention are useful for producing plants
CC   which are resistant and/or tolerant to a herbicide comprising fomesafen
CC   and/or acifluorfen. Methods of the invention are useful for providing
CC   plants with further desired agronomic trait, especially resistant to a
CC   herbicide, comprising glyphosate or its salt. Further desired agronomic
CC   traits include insect resistance, nematode resistance, stress tolerance,
CC   altered field, altered nutritional value, altered quality or any other
CC   desirable agronomic trait. GST or its variant is also useful as a
CC   selectable marker gene. The present sequence represents a Glycine max
CC   clone SE3.03B09 sequence. The GST 3.3 protein sequence of the invention
CC   is specifically claimed to be not a part of the present sequence
XX
SQ   Sequence 234 AA;

  Query Match             83.7%;  Score 1000.5;  DB 4;  Length 234;
  Best Local Similarity   84.6%;  Pred. No. 1.1e-97;
  Matches  192;  Conservative   17;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ||:||| |||||||||||||||||:||||||||||:||||||||||||||||||||
Db          6 DLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSDLLLKSNPVHKKIPVFFHGDKV 65

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||||||||||| |: | |||||||||||||||| :|||:|| |||:|||  ||||||
Db         66 ICESAIIVEYIDEAWT-NVPSILPQNAYDRANARFWFAYIDEKWFTSLRSVLVAEDDEAK 124

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              | ||||||| ||::|||||| ||||||||||:|||:|||||||||::|| | |: |||||
Db        125 KPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFIDIGFGSFLSWMRVIEEMSGRKLLD 184

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAAA 231
              | |:|||| |||||||||||||:||||:|||||||||||||| |||||
Db        185 EKKHPGLTQWAETFAADPAVKGILPETDKLVEFAKILQLKWTAAAAA 231


RESULT 7
ABB90937
ID   ABB90937 standard; protein; 227 AA.
XX
AC   ABB90937;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 148.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP009892.
XX
PR   28-AUG-2001; 2001WO-EP009892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms.
XX
PS   Claim 5; SEQ ID NO 148; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins (ABB90790-ABB94016)
CC   for herbicidally active compounds, comprising aligning and comparing
CC   nucleic acid or amino acid sequences from plant with nucleic acid or
CC   amino acid sequences from non-plant organisms using suitable search
CC   parameters, where plant sequences having an E-value greater by a factor
CC   of 3 than the E-value of most similar non-plant sequences are selected.
CC   The polypeptides or nucleic acids encoding them are useful for
CC   identifying modulators. The identified modulators are useful as
CC   herbicides
XX
SQ   Sequence 227 AA;

  Query Match             53.4%;  Score 638;  DB 5;  Length 227;
```

```
  Best Local Similarity   54.3%;  Pred. No. 3.7e-59;
  Matches  120;  Conservative   34;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          1 MAERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFH 60
              ||  |::|:|:| | : :| :|||:|| : || ::||   |||||||||||||:||  |
Db          1 MATEDVKLIGSWASVYVMRARIALHLKSISYEFLQETYGSKSELLLKSNPVHKKMPVLIH 60

Qy         61 GDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATED 120
               || :||| ||| |||| |::   |||| : |||| |||| :||||:|  |::|:|  :
Db         61 ADKPVCESNIIVHYIDEAWNSSGPSILPSHPYDRAIARFWAAYIDDQWFISVRSILTAQG 120

Qy        121 DEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNER 180
              || ||    | ||  : :|: || ||:|| :| || ||::||  |||| : || |
Db        121 DEEKKAAIAQVEERTKLLEKAFNDCSQGKPFFNGDHIGYLDIALGSFLGWWRVVELDANH 180

Qy        181 KLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKIL 221
              | ||||| | |  ||| |  ||||| ::||  || |||: |
Db        181 KFLDETKTPSLVKWAERFCDDPAVKPIMPEITKLAEFARKL 221


RESULT 8
AAB47340
ID   AAB47340 standard; protein; 233 AA.
XX
AC   AAB47340;
XX
DT   19-SEP-2001  (first entry)
XX
DE   GST3.
XX
KW   Glutathione-S-transferase; GST; opium poppy; conjugation; herbicide;
KW   electrophilic compound; glutathione; GSH; detoxification; xenobiotic;
KW   transgenic plant; stress resistance; pathogen; grazing pest.
XX
OS   Pavaver somniferum.
XX
PN   WO200153501-A2.
XX
PD   26-JUL-2001.
XX
PF   18-JAN-2001; 2001WO-IB000205.
XX
PR   18-JAN-2000; 2000US-0176708P.
XX
PA   (FACC/) FACCHINI P J.
XX
PI   Facchini PJ;
XX
DR   WPI; 2001-457612/49.
DR   N-PSDB; AAC86194.
XX
PT   Glutathione-s-transferase polypeptide(s) useful for producing transgenic
PT   plants comprising resistance to biotic and abiotic stress.
XX
PS   Claim 1; Page 10; 77pp; English.
XX
CC   This sequence shows a glutathione-S-transferase (GST) polypeptide, GST3,
CC   from the opium poppy. GST's catalyse the conjugation of electrophilic
CC   compounds to glutathione (GSH). They have been shown to be involved in
CC   detoxification of xenobiotics. The novel GST's of the invention GST1,
CC   GST2 and GST3, show extensive homology with tau type GST's from a variety
CC   of plant species, especially within the N-terminal region. This domain
CC   has been shown to be responsible for recognition of GSH and the ability
CC   of the proteins to form dimers. The C-terminal domain is responsible for
CC   substrate specificity and varies greatly between other GST's. The GST's
CC   of the invention exhibit strong glutathione conjugating activity towards
CC   a model substrate, 1-chloro-2,4-dinitrobenzene that could be inhibited in
CC   the presence of hydroxycinnamic acid amines of tyramine. The expression
CC   of the GST polypeptide in a transgenic plant is useful to produce a plant
CC   which has increased stress resistance. The plant has increased resistance
CC   to pathogens, herbicides or grazing pests. The GST polypeptide is useful
CC   to identify binding agents inhibitors and substrates
XX

SQ   Sequence 233 AA;

  Query Match             52.2%;  Score 624;  DB 4;  Length 233;
  Best Local Similarity   54.0%;  Pred. No. 1.2e-57;
  Matches  116;  Conservative   39;  Mismatches   60;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::|||||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKIPVMIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|| : ||::|| || :: :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGEEKIGYIDIAFGCYIGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              ||| |||| ||| | ||  || ::|||: |:|||
Db        188 ETKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 9
AAB47339
ID   AAB47339 standard; protein; 233 AA.
XX
AC   AAB47339;
XX
DT   19-SEP-2001  (first entry)
XX
DE   GST2.
XX
KW   Glutathione-S-transferase; GST; opium poppy; conjugation; herbicide;
KW   electrophilic compound; glutathione; GSH; detoxification; xenobiotic;
KW   transgenic plant; stress resistance; pathogen; grazing pest.
XX
OS   Pavaver somniferum.
XX
PN   WO200153501-A2.
XX
PD   26-JUL-2001.
XX
PF   18-JAN-2001; 2001WO-IB000205.
XX
PR   18-JAN-2000; 2000US-0176708P.
XX
PA   (FACC/) FACCHINI P J.
XX
PI   Facchini PJ;
XX
DR   WPI; 2001-457612/49.
DR   N-PSDB; AAC86193.
XX
PT   Glutathione-s-transferase polypeptide(s) useful for producing transgenic
PT   plants comprising resistance to biotic and abiotic stress.
XX
PS   Claim 1; Page 9; 77pp; English.
XX
CC   This sequence shows a glutathione-S-transferase (GST) polypeptide, GST2,
CC   from the opium poppy. GST's catalyse the conjugation of electrophilic
CC   compounds to glutathione (GSH). They have been shown to be involved in
CC   detoxification of xenobiotics. The novel GST's of the invention GST1,
CC   GST2 and GST3, show extensive homology with tau type GST's from a variety
CC   of plant species, especially within the N-terminal region. This domain
CC   has been shown to be responsible for recognition of GSH and the ability
CC   of the proteins to form dimers. The C-terminal domain is responsible for
CC   substrate specificity and varies greatly between other GST's. The GST's
CC   of the invention exhibit strong glutathione conjugating activity towards
CC   a model substrate, 1-chloro-2,4-dinitrobenzene that could be inhibited in
CC   the presence of hydroxycinnamic acid amines of tyramine. The expression
CC   of the GST polypeptide in a transgenic plant is useful to produce a plant
CC   which has increased stress resistance. The plant has increased resistance
```

```
CC   to pathogens, herbicides or grazing pests. The GST polypeptide is useful
CC   to identify binding agents inhibitors and substrates
XX
SQ   Sequence 233 AA;

  Query Match            52.1%;  Score 623;  DB 4;  Length 233;
  Best Local Similarity  54.0%;  Pred. No. 1.5e-57;
  Matches  116;  Conservative   39;  Mismatches   60;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::||:||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKMPVLIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|||: ||:||| || :: :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGGEKIGYVDIAFGCYVGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              | | |||| ||| | ||  || ::|||: |:|||
Db        188 EEKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 10
AAB47341
ID   AAB47341 standard; protein; 233 AA.
XX
AC   AAB47341;
XX
DT   19-SEP-2001  (first entry)
XX
DE   GST consensus sequence.
XX
KW   Glutathione-S-transferase; GST; opium poppy; conjugation; herbicide;
KW   electrophilic compound; glutathione; GSH; detoxification; xenobiotic;
KW   transgenic plant; stress resistance; pathogen; grazing pest.
XX
OS   Pavaver somniferum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 58
FT                   /label= Met, Ile
FT   Misc-difference 61
FT                   /label= Met, Leu
FT   Misc-difference 158
FT                   /label= Gly, Glu
FT   Misc-difference 163
FT                   /label= Val, Ile
FT   Misc-difference 171
FT                   /label= Val, Ile
FT   Misc-difference 189
FT                   /label= Glu, Thr
XX
PN   WO200153501-A2.
XX
PD   26-JUL-2001.
XX
PF   18-JAN-2001; 2001WO-IB000205.
XX
PR   18-JAN-2000; 2000US-0176708P.
XX
PA   (FACC/) FACCHINI P J.
XX
PI   Facchini PJ;
XX
DR   WPI; 2001-457612/49.
XX
PT   Glutathione-s-transferase polypeptide(s) useful for producing transgenic
PT   plants comprising resistance to biotic and abiotic stress.
XX
PS   Claim 1; Page 12; 77pp; English.
XX
CC   This sequence shows a consenus glutathione-S-transferase (GST)
CC   polypeptide from the opium poppy. GST's catalyse the conjugation of
CC   electrophilic compounds to glutathione (GSH). They have been shown to be
CC   involved in detoxification of xenobiotics. The novel GST's of the
CC   invention GST1, GST2 and GST3, show extensive homology with tau type
CC   GST's from a variety of plant species, especially within the N-terminal
CC   region. This domain has been shown to be responsible for recognition of
CC   GSH and the ability of the proteins to form dimers. The C-terminal domain
CC   is responsible for substrate specificity and varies greatly between other
CC   GST's. The GST's of the invention exhibit strong glutathione conjugating
CC   activity towards a model substrate, 1-chloro-2,4-dinitrobenzene that
CC   could be inhibited in the presence of hydroxycinnamic acid amines of
CC   tyramine. The expression of the GST polypeptide in a transgenic plant is
CC   useful to produce a plant which has increased stress resistance. The
CC   plant has increased resistance to pathogens, herbicides or grazing pests.
CC   The GST polypeptide is useful to identify binding agents inhibitors and
CC   substrates
XX
SQ   Sequence 233 AA;

  Query Match            50.8%;  Score 607;  DB 4;  Length 233;
  Best Local Similarity  53.0%;  Pred. No. 7.7e-56;
  Matches  114;  Conservative   37;  Mismatches   64;  Indels    0;  Gaps    0;

Qy          5 DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETLNPKSELLLKSNPVHKKIPVFFHGDKV 64
              ::::|| | ||| :| :||||:| : | ::|||   ||||||||||::|| ||  ||||
Db          8 EVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSKSELLLKSNPIYKKKPVXIHGDKP 67

Qy         65 ICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDDEAK 124
              |||| |||:|||:||::   ||:| : || : |||| :|||||:  ||  :  ::| | |
Db         68 ICESMIIVQYIDDVWASAGHSIIPSDPYDASIARFWATYIDDKFFPSLMGIAKSKDAEEK 127

Qy        125 KLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERKLLD 184
              |   |||      :|| : | |:|| :|| : ||: || || :  :|||:| ||  || |
Db        128 KAAIEQAIAAFGILEEAYQKTSKGKDFFGXEKIGYXDIAFGCYXGWIRVTEKMNGIKLFD 187

Qy        185 ETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAK 219
              | | |||| ||| | ||  || ::|||: |:||||
Db        188 EXKVPGLTKWAEKFCADETVKSVMPETDALMEFAK 222


RESULT 11
ABB91390
ID   ABB91390 standard; protein; 234 AA.
XX
AC   ABB91390;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 601.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP009892.
XX
PR   28-AUG-2001; 2001WO-EP009892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
```

```
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms.
XX
PS   Claim 5; SEQ ID NO 601; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins (ABB90790-ABB94016)
CC   for herbicidally active compounds, comprising aligning and comparing
CC   nucleic acid or amino acid sequences from plant with nucleic acid or
CC   amino acid sequences from non-plant organisms using suitable search
CC   parameters, where plant sequences having an E-value greater by a factor
CC   of 3 than the E-value of most similar non-plant sequences are selected.
CC   The polypeptides or nucleic acids encoding them are useful for
CC   identifying modulators. The identified modulators are useful as
CC   herbicides
XX
SQ   Sequence 234 AA;

  Query Match             49.9%;  Score 596;  DB 5;  Length 234;
  Best Local Similarity   49.6%;  Pred. No. 1.1e-54;
  Matches  115;  Conservative   46;  Mismatches   69;  Indels    2;  Gaps    2;

Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | |: :::||| |:||:|:| :||| || :||: ||| |   |||||||||||||:||
Db          1 MGEKEEVKLLGVWYSPYAIRPKIALRLKSVDYDYVEENLFGSKSELLLKSNPVHKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
               | :| | ||  ||||||| |:::| |||| : |||| ||||  ::|:||  :|:    |
Db         61 LHNNKPIVESLNIVEYIDETWNSSAPSILPSHPYDRALARFWSDFVDNKWFPALRMAAIT 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
             : ::||    |: || | ::|: |   |:|| :|||: |||:|| |||:  ::  |
Db        121 KSEDAKAKAMEEVEEGLLQLEDAFVSISKGKPFFGGEAIGFMDICFGSFVVLLKAREKFK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              |||||:| | |  ||: | :|  || : || ||: || : |:::  :||:
Db        181 AEKLLDESKTPSLCKWADRFLSDETVKNVAPEIEKVAEFLQELEVRAQSAAS 232


RESULT 12
AAG26676
ID   AAG26676 standard; protein; 234 AA.
XX
AC   AAG26676;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 31221.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
```

PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 49.0%; Score 585.5; DB 3; Length 234;
Best Local Similarity 48.5%; Pred. No. 1.5e-53;
Matches 111; Conservative 48; Mismatches 69; Indels 1; Gaps 1;

```
Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVFFHG 61
              :::::||| |:||:|:| :||| || :||: ||| |   |||||||||||||:||  |
Db          4 KQEVKLLGVWYSPYAIRTKIALRLKSVDYDYVEEDLFGSKSELLLKSNPVHKKVPVLLHY 63

Qy         62 DKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDD 121
              :| | ||  ||||||| |:::| |||| : |||| ||||  ::|:||  :|:      : :
Db         64 NKPIVESLNIVEYIDETWNSSAPSILPSHPYDRALARFWSDFVDNKWFPALRVAAIAKSE 123

Qy        122 EAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERK 181
              :||    |: || | ::|: |   |:|| :|||: |||:|| ||||:  ::  |     |
Db        124 DAKAKAMEEVEEGLLQLEDAFVSISKGKPFFGGEVIGFMDICFGSFVVLLKAREKFKAEK 183

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              ||||:| | |  ||: | :|  || : |: ||: || : |:::  :||:
Db        184 LLDESKTPSLCKWADRFLSDETVKNVAPDIEKVAEFLQELEVRAQSAAS 232
```

RESULT 13
AAG26675
ID AAG26675 standard; protein; 281 AA.
XX
AC AAG26675;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 31220.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P:
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 49.0%; Score 585.5; DB 3; Length 281;
Best Local Similarity 48.5%; Pred. No. 1.9e-53;
Matches 111; Conservative 48; Mismatches 69; Indels 1; Gaps 1;

```
Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVFFHG 61
              :::::||| |:||:|:| :||| || :||: ||| |   ||||||||||||||:||  |
Db         51 KQEVKLLGVWYSPYAIRTKIALRLKSVDYDYVEEDLFGSKSELLLKSNPVHKKVPVLLHY 110

Qy         62 DKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDD 121
              :| | ||  |||||| |:::| |||| : |||| ||||  ::|:||  :|:     : :
Db        111 NKPIVESLNIVEYIDETWNSSAPSILPSHPYDRALARFWSDFVDNKWFPALRVAAIAKSE 170

Qy        122 EAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERK 181
              :||    |: || | ::|: |   |:|| :|||: |||:|| ||||:  ::  |     |
Db        171 DAKAKAMEEVEEGLLQLEDAFVSISKGKPFFGGEVIGFMDICFGSFVVLLKAREKFKAEK 230

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              ||||:| | |  ||: | :|  || : |: ||: || : |:::  :||:
Db        231 LLDESKTPSLCKWADRFLSDETVKNVAPDIEKVAEFLQELEVRAQSAAS 279
```

RESULT 14
AAG26674.
ID AAG26674 standard; protein; 287 AA.
XX
AC AAG26674;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 31219.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 49.0%; Score 585.5; DB 3; Length 287;
Best Local Similarity 48.5%; Pred. No. 2e-53;
Matches 111; Conservative 48; Mismatches 69; Indels 1; Gaps 1;

```
Qy          3 ERDLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVFFHG 61
              :::::||| |:||:|:| :||| || :||: ||| |   |||||||||||||:||  |
Db         57 KQEVKLLGVWYSPYAIRTKIALRLKSVDYDYVEEDLFGSKSELLLKSNPVHKKVPVLLHY 116

Qy         62 DKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLATEDD 121
              :| | ||  |||||| |:::| |||| : |||| ||||  ::|:||  :|:     : :
Db        117 NKPIVESLNIVEYIDETWNSSAPSILPSHPYDRALARFWSDFVDNKWFPALRVAAIAKSE 176

Qy        122 EAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMNERK 181
              :||    |: || | ::|: |   |:|| :||: |||:|| ||||:  ::  |     |
Db        177 DAKAKAMEEVEEGLLQLEDAFVSISKGKPFFGGEVIGFMDICFGSFVVLLKAREKFKAEK 236

Qy        182 LLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEFAKILQLKWAAAAA 230
              ||||:| | |  ||: | :|  || : |: ||: || : |:::  :||:
Db        237 LLDESKTPSLCKWADRFLSDETVKNVAPDIEKVAEFLQELEVRAQSAAS 285
```

RESULT 15
AAG53029
ID AAG53029 standard; protein; 233 AA.
XX
AC AAG53029;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67473.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.

PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.7%; Score 558; DB 3; Length 233;
Best Local Similarity 48.4%; Pred. No. 1.2e-50;
Matches 106; Conservative 42; Mismatches 69; Indels 2; Gaps 2;

```
Qy          1 MAER-DLRLLGAWFSPFALRVQIALNLKGLDYEVVEETL-NPKSELLLKSNPVHKKIPVF 58
              | || :::||| |:||  :| :||| || :||: ||| |   ||||||||||: ||:||
Db          1 MGEREEVKLLGTWYSPVVIRAKIALRLKSVDYDYVEEDLFGSKSELLLKSNPIFKKVPVL 60

Qy         59 FHGDKVICESAIIVEYIDEVWSNNALSILPQNAYDRANARFWVSYIDDKWLTSLKSVLAT 118
               |  | :| |  ||||||| |:::  |||| : |||| ||||  ::||||| :| : :
Db         61 IHNTKPVCVSLNIVEYIDETWNSSGSSILPSHPYDRALARFWSVFVDDKWLPTLMAAVVA 120

Qy        119 EDDEAKKLHFEQAEEVLEKVEEVFNKCSEGKAYFGGDTIGFVDIGFGSFLSFIRVSENMN 178
              : :|||    |: || | ::|  |   |:||::|||:||||:||  ||||  ::  | :
Db        121 KSEEAKAKGMEEVEEGLLQLEAAFIALSKGKSFFGGETIGFIDICLGSFLVLLKAREKLK 180

Qy        179 ERKLLDETKYPGLTLWAETFAADPAVKGLLPETEKLVEF 217
              |:||| | | |  ||  | ::  || ::|: :|: :|
Db        181 NEKILDELKTPSLYRWANQFLSNEMVKNVVPDIDKVAKF 219
```

Search completed: October 8, 2004, 09:21:42
Job time : 127 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 12, 2004, 23:48:11 ; Search time 3928 Seconds
(without alignments)
10129.558 Million cell updates/sec

Title: US-10-088-945A-14
Perfect score: 918
Sequence: 1 cataaaactccacatttcct.....aatatagtaaaaaaaaaaaa 918

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	918	100.0	918	6	AX100331	AX100331 Sequence
2	620.2	67.6	895	6	AR123130	AR123130 Sequence
3	620.2	67.6	895	6	AR124082	AR124082 Sequence
4	620.2	67.6	895	6	AX100355	AX100355 Sequence
5	620.2	67.6	895	8	AF243360	AF243360 Glycine m
6	288	31.4	896	6	AX203249	AX203249 Sequence
7	288	31.4	896	8	AF118926	AF118926 Papaver s
8	270	29.4	998	6	AX203245	AX203245 Sequence
9	270	29.4	998	8	AF118924	AF118924 Papaver s
10	268.4	29.2	947	6	AX203247	AX203247 Sequence
11	268.4	29.2	947	8	AF118925	AF118925 Papaver s
12	265.4	28.9	962	8	AY091102	AY091102 Arabidops
13	264.8	28.8	988	8	AF159229	AF159229 Gossypium
14	264.4	28.8	684	8	AF288191	AF288191 Arabidops
15	263.8	28.7	964	8	AB039930	AB039930 Arabidops
16	241.6	26.3	715	8	BT000940	BT000940 Arabidops
17	240	26.1	684	6	AX507981	AX507981 Sequence
18	240	26.1	684	8	AF288190	AF288190 Arabidops
19	232.2	25.3	705	6	AX507963	AX507963 Sequence
20	231.4	25.2	938	8	AF370480	AF370480 Arabidops
21	204.2	22.2	855	8	AF177944	AF177944 Petroseli
22	197.4	21.5	866	8	AY084992	AY084992 Arabidops
23	193.6	21.1	513	6	AX506970	AX506970 Sequence
24	185.6	20.2	624	8	AF288192	AF288192 Arabidops
25	184.4	20.1	593	8	AF064201	AF064201 Gossypium
26	184	20.0	810	6	AX653837	AX653837 Sequence
27	182.4	19.9	1151	8	AK107435	AK107435 Oryza sat
28	181.8	19.8	684	6	AX506000	AX506000 Sequence
29	181.8	19.8	684	8	AF288193	AF288193 Arabidops
30	181.8	19.8	684	8	AY094051	AY094051 Arabidops
31	181.8	19.8	754	8	AY044324	AY044324 Arabidops
32	181.8	19.8	901	8	AY050343	AY050343 Arabidops
33	181.2	19.7	702	6	AX653944	AX653944 Sequence
34	181.2	19.7	928	8	AK073944	AK073944 Oryza sat
35	181.2	19.7	929	8	AK061304	AK061304 Oryza sat
36	181.2	19.7	943	8	AY271620	AY271620 Oryza sat
37	181.2	19.7	947	8	AF309379	AF309379 Oryza sat
38	179	19.5	705	8	BT006220	BT006220 Arabidops
39	179	19.5	927	8	AK119143	AK119143 Arabidops
40	178.6	19.5	1040	6	AX100349	AX100349 Sequence
41	178.6	19.5	1040	8	AF051214	AF051214 Picea mar
42	178.4	19.4	696	6	AX652944	AX652944 Sequence
43	177.4	19.3	1134	8	AK059726	AK059726 Oryza sat
44	177.4	19.3	1223	8	AF402801	AF402801 Oryza sat
45	175.4	19.1	902	6	AX100350	AX100350 Sequence

ALIGNMENTS

RESULT 1
AX100331
LOCUS AX100331 918 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 14 from Patent WO0121770.
ACCESSION AX100331
VERSION AX100331.1 GI:13619349
KEYWORDS .
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1
AUTHORS Andrews,C.J., Jepson,I., Townson,J.K., Edwards,R., Cummins,I.D. and

```
          Skipsey,M.D.
  TITLE   Gst sequences from soybean and their use in the production of
          herbicide resistant plants
  JOURNAL Patent: WO 0121770-A 14 29-MAR-2001;
          ZENECA LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Glycine max"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3847"
ORIGIN

  Query Match            100.0%;  Score 918;  DB 6;  Length 918;
  Best Local Similarity  100.0%;  Pred. No. 8.4e-205;
  Matches  918;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATAAAACTCCACATTTCCTGCTGAGTAACCTAACAAAACAAACACAATATTGCTCCGTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATAAAACTCCACATTTCCTGCTGAGTAACCTAACAAAACAAACACAATATTGCTCCGTG 60

Qy         61 TTTGACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTGACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGG 120

Qy        121 TTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 180

Qy        181 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240

Qy        241 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 300

Qy        301 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 360

Qy        361 GATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 420

Qy        421 AAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAA 480

Qy        481 GAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTC 540

Qy        541 GGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGA 600

Qy        601 GTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACC 660

Qy        661 CTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAA 720

Qy        721 AAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAG 780

Qy        781 TAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATA 840

Qy        841 TCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGATAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGATAA 900

Qy        901 TATAGTAAAAAAAAAAAA 918
              ||||||||||||||||||
Db        901 TATAGTAAAAAAAAAAAA 918


RESULT 2
AR123130
LOCUS       AR123130                 895 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 7 from patent US 6168954.
ACCESSION   AR123130
VERSION     AR123130.1  GI:14108096
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   McGonigle,B. and O'Keefe,D.P.
  TITLE     Soybean glutathione-S-transferase enzymes
  JOURNAL   Patent: US 6168954-A 7 02-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match            67.6%;  Score 620.2;  DB 6;  Length 895;
  Best Local Similarity  88.0%;  Pred. No. 6.4e-135;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| |||||||| | ||||||| |||||||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || |||||||||||||||||||| ||||||||||||||||  ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||| |||||||| |||||||| |||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              |||||  |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              ||||  |||||||||||||| |||||||  |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | |||||||||||||||||||||||||||||  ||||||||  ||||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||||||| ||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
```

```
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810

Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870

Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913

Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 3
AR124082
LOCUS       AR124082                 895 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 7 from patent US 6171839.
ACCESSION   AR124082
VERSION     AR124082.1  GI:14109443
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   McGonigle,B. and O'Keefe,D.P.
  TITLE     Soybean glutathione-S-transferase enzymes
  JOURNAL   Patent: US 6171839-A 7 09-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           67.6%;  Score 620.2;  DB 6;  Length 895;
  Best Local Similarity  88.0%;  Pred. No. 6.4e-135;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156

Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216

Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276

Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336

Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396

Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456

Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516

Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576

Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636

Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696

Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756

Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810

Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870

Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913

Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 4
AX100355
LOCUS       AX100355                 895 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 38 from Patent WO0121770.
ACCESSION   AX100355
VERSION     AX100355.1  GI:13619371
KEYWORDS    .
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1
  AUTHORS   Andrews,C.J., Jepson,I., Townson,J.K., Edwards,R., Cummins,I.D. and
            Skipsey,M.D.
  TITLE     Gst sequences from soybean and their use in the production of
            herbicide resistant plants
  JOURNAL   Patent: WO 0121770-A 38 29-MAR-2001;
            ZENECA LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="Glycine max"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3847"
ORIGIN

  Query Match           67.6%;  Score 620.2;  DB 6;  Length 895;
  Best Local Similarity  88.0%;  Pred. No. 6.4e-135;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156

Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216

Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276

Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336

Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
```

```
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | ||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| ||||||| || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| ||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||||| |||||||||||||||| ||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || ||||||||||||||||||||||||||||||  || | |
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  |||||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |      | |   | || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 5
AF243360
LOCUS       AF243360                 895 bp    mRNA    linear   PLN 28-NOV-2000
DEFINITION  Glycine max glutathione S-transferase GST 5 mRNA, complete cds.
ACCESSION   AF243360
VERSION     AF243360.1  GI:11385424
KEYWORDS    .
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   McGonigle,B., Keeler,S.J., Lau,S.M., Koeppe,M.K. and O'Keefe,D.P.
  TITLE     A genomics approach to the comprehensive analysis of the
            glutathione S-transferase gene family in soybean and maize
  JOURNAL   Plant Physiol. 124 (3), 1105-1120 (2000)
  MEDLINE   20532786
   PUBMED   11080288
REFERENCE   2  (bases 1 to 895)
  AUTHORS   McGonigle,B. and O'Keefe,D.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) AgBiotechnology, DuPont, PO Box 80402,
            Wilmington, DE 19880-0402, USA
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
     CDS             36. .740
                     /EC_number="2.5.1.18"
                     /codon_start=1
                     /product="glutathione S-transferase GST 5"
                     /protein_id="AAG34795.1"
                     /db_xref="GI:11385425"
                     /translation="MSKSEDLKLLGGWFSPFALRVQIALNLKGLEYEVVEETLNPKSD
                     LLLKSNPVHKKIPVFFHGDKVICESAIIVEYIDEAWTNVPSILPQNAYDRANARFWFA
                     YIDEKWFTSLRSVLVAEDDEAKKPHFEQAEEGLERLEEVFNKYSEGKAYFGGDSIGFI
                     DIGFGSFLSWMRVIEEMSGRKLLDEKKHPGLTQWAETFAADPAVKGILPETDKLVEFA
                     KILQLKWTAAAAAAAK"
ORIGIN

  Query Match           67.6%;  Score 620.2;  DB 8;  Length 895;
  Best Local Similarity  88.0%;  Pred. No. 6.4e-135;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| ||||||||| |  |||||||| ||||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || |||||||||||||||||| ||||||||||||||||  ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||||| |||||||| |||||||| |||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | |||||||||||||||||||  .
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | ||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| ||||||| || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||||| |||||||||||||||| ||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || ||||||||||||||||||||||||||||||  || | |
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  |||||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |      | |   | || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 6
AX203249
LOCUS       AX203249                 896 bp    DNA     linear   PAT 30-AUG-2001
```

```
DEFINITION  Sequence 5 from Patent WO0153501.
ACCESSION   AX203249
VERSION     AX203249.1  GI:15392619
KEYWORDS    .
SOURCE      Papaver somniferum (opium poppy)
  ORGANISM  Papaver somniferum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
            Papaveraceae; Papaver.
REFERENCE   1
  AUTHORS   Facchini,P.J.
  TITLE     Glutathione-s-transferase nucleic acids and polypeptides and
            methods of use thereof
  JOURNAL   Patent: WO 0153501-A 5 26-JUL-2001;
            Facchini, Peter James (CA)
FEATURES             Location/Qualifiers
     source          1. .896
                     /organism="Papaver somniferum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3469"
ORIGIN

  Query Match           31.4%;  Score 288;  DB 6;  Length 896;
  Best Local Similarity   59.6%;  Pred. No. 5.7e-57;
  Matches  486;  Conservative    0;  Mismatches  330;  Indels    0;  Gaps    0;

Qy         96 GGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCT 155
              |||  | | | |||| ||||  |||   ||||||||||    |||||      ||||| ||
Db         81 GGAGGTAAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACT 140

Qy        156 TAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATT 215
               ||| | ||     |  | ||| |  || ||||||||| ||   |  |||||||||| |
Db        141 CAACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACT 200

Qy        216 GCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGT 275
              |||||  || || || ||| |  |||||||||| || ||| |  | ||||| ||||||
Db        201 TCTTCTGAAATCAAATCCTATTTACAAGAAAATTCCTGTTATGATTCATGGTGATAAACC 260

Qy        276 CATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCT 335
              ||| ||||||||     ||||| ||| ||||||| ||||| ||||||  |      ||  |
Db        261 CATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTTTGGGCTTCTGCTGGACA 320

Qy        336 CTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACAT 395
              || ||| | ||    || | ||||||    | | ||| || || ||||    | |||||
Db        321 TTCTATCATCCCGTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACAT 380

Qy        396 CGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAA 455
              ||||||||||    || |||| || |   |  ||  |  || | | ||||  || | |||
Db        381 TGATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAA 440

Qy        456 GAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAA 515
              ||    |  | |||| || |||   | ||   |||  |   |||||||||  |    | ||
Db        441 AAAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTGGAAGAAGCTTATCAGAA 500

Qy        516 GTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTT 575
             ||| |||| || | || |||||| | ||| |  |||||||   |||| ||||   ||
Db        501 AACTAGTAAAGGAAAAGACTTTTTCGGGGAAGAAAAAATTGGATACATTGATATTGCATT 560

Qy        576 TGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGA 635
              |||  | | | |  |||  ||||||||  |||||||  |||||||  ||  |||  ||||
Db        561 TGGGTGTTATATAGGTTGGATTAGAGTTACAGAGAAAATGAATGGAATCAAACTATTTGA 620

Qy        636 TGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGT 695
              ||||||  ||    ||  ||  |  ||  |||||||||  |  |||  ||| |||    | ||
Db        621 TGAAACAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACAGT 680

Qy        696 GAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAA 755
              ||      || |||| || |||||     ||  | ||||||||  ||||    |       |
Db        681 TAAATCTGTTATGCCTGAAACTGATGCTCTCATGGAGTTTGCTAAGAAGATCTTTGGATC 740

Qy        756 ATGGGCTGCTGCAGCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAATTTCA 815
              | || || | |          ||  |  ||  | |||   |||  ||    ||  || |
Db        741 TAAGCCTCCTCCTTCAAACTAGAAAAAGTTGTTAACAATGAAATATCTTAGAGATGTTTA 800

Qy        816 AAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATATATTTA 875
              |    ||||  |      ||        | ||  || ||| ||          ||| || |  |
Db        801 AGCTTTGTGTTTGTTTTTTCAGTGTTGTGTAGCAATGCTTAAGAACTGTTTGTAGAAATG 860

Qy        876 AAAGTCAAAATAAATGTTATGATAATATAGTAAAAA 911
              |      || |  |     ||  | || | |  |||||
Db        861 ATCAAGAACAGTAGCTGTAAAAAAAAAAAAAAAAAA 896


RESULT 7
AF118926
LOCUS       AF118926                896 bp    mRNA    linear   PLN 02-JAN-2000
DEFINITION  Papaver somniferum glutathione S-transferase 3 mRNA, complete cds.
ACCESSION   AF118926
VERSION     AF118926.1  GI:6652873
KEYWORDS    .
SOURCE      Papaver somniferum (opium poppy)
  ORGANISM  Papaver somniferum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
            Papaveraceae; Papaver.
REFERENCE   1  (bases 1 to 896)
  AUTHORS   Facchini,P.J. and Yu,M.
  TITLE     Molecular cloning and characterization of a glutathione
            S-transferase gene family from opium poppy
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 896)
  AUTHORS   Facchini,P.J. and Yu,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1999) Biological Sciences, University of Calgary,
            2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada
FEATURES             Location/Qualifiers
     source          1. .896
                     /organism="Papaver somniferum"
                     /mol_type="mRNA"
                     /cultivar="Marianne"
                     /db_xref="taxon:3469"
     CDS             61. .762
                     /EC_number="2.5.1.18"
                     /note="GST3"
                     /codon_start=1
                     /product="glutathione S-transferase 3"
                     /protein_id="AAF22519.1"
                     /db_xref="GI:6652874"
                     /translation="MAGSGSEEVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSK
                     SELLLKSNPIYKKIPVMIHGDKPICESMIIVQYIDDVWASAGHSIIPSDPYDASIARF
                     WATYIDDKFFPSLMGIAKSKDAEEKKAAIEQAIAAFGILEEAYQKTSKGKDFFGEEKI
                     GYIDIAFGCYIGWIRVTEKMNGIKLFDETKVPGLTKWAEKFCADETVKSVMPETDALM
                     EFAKKIFGSKPPPSN"
ORIGIN

  Query Match           31.4%;  Score 288;  DB 8;  Length 896;
  Best Local Similarity   59.6%;  Pred. No. 5.7e-57;
  Matches  486;  Conservative    0;  Mismatches  330;  Indels    0;  Gaps    0;

Qy         96 GGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCT 155
              |||  | | | |||| ||||  |||   ||||||||||    |||||      ||||| ||
Db         81 GGAGGTAAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACT 140

Qy        156 TAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATT 215
               ||| | ||     |  | ||| |  || ||||||||| ||   |  |||||||||| |
Db        141 CAACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACT 200

Qy        216 GCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGT 275
              |||||  || || || ||| |  |||||||||| || ||| |  | ||||| ||||||
Db        201 TCTTCTGAAATCAAATCCTATTTACAAGAAAATTCCTGTTATGATTCATGGTGATAAACC 260

Qy        276 CATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCT 335
              ||| ||||||||     ||||| ||| ||||||| ||||| ||||||  |      ||  |
Db        261 CATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTTTGGGCTTCTGCTGGACA 320

Qy        336 CTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACAT 395
```

```
Db        321 TTCTATCATCCCGTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACAT 380

Qy        396 CGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAA 455

Db        381 TGATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAA 440

Qy        456 GAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAA 515

Db        441 AAAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTGGAAGAAGCTTATCAGAA 500

Qy        516 GTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTT 575

Db        501 AACTAGTAAAGGAAAAGACTTTTTCGGGGAAGAAAAAATTGGATACATTGATATTGCATT 560

Qy        576 TGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGA 635

Db        561 TGGGTGTTATATAGGTTGGATTAGAGTTACAGAGAAAATGAATGGAATCAAACTATTTGA 620

Qy        636 TGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTGCTGCTGATCCTGCTGT 695

Db        621 TGAAACAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACAGT 680

Qy        696 GAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAA 755

Db        681 TAAATCTGTTATGCCTGAAACTGATGCTCTCATGGAGTTTGCTAAGAAGATCTTTGGATC 740

Qy        756 ATGGGCTGCTGCAGCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAATTTCA 815

Db        741 TAAGCCTCCTCCTTCAAACTAGAAAAAGTTGTTAACAATGAAATATCTTAGAGATGTTTA 800

Qy        816 AAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATATATTTA 875

Db        801 AGCTTTGTGTTTGTTTTTTCAGTGTTGTGTAGCAATGCTTAAGAACTGTTTGTAGAAATG 860

Qy        876 AAAGTCAAAATAAATGTTATGATAATATAGTAAAAA 911

Db        861 ATCAAGAACAGTAGCTGTAAAAAAAAAAAAAAAAA 896


RESULT 8
AX203245
LOCUS       AX203245                998 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 1 from Patent WO0153501.
ACCESSION   AX203245
VERSION     AX203245.1  GI:15392617
KEYWORDS    .
SOURCE      Papaver somniferum (opium poppy)
  ORGANISM  Papaver somniferum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
            Papaveraceae; Papaver.
REFERENCE   1
  AUTHORS   Facchini,P.J.
  TITLE     Glutathione-s-transferase nucleic acids and polypeptides and
            methods of use thereof
  JOURNAL   Patent: WO 0153501-A 1 26-JUL-2001;
            Facchini, Peter James (CA)
FEATURES             Location/Qualifiers
     source          1. .998
                     /organism="Papaver somniferum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3469"
ORIGIN

  Query Match            29.4%;  Score 270;  DB 6;  Length 998;
  Best Local Similarity  63.6%;  Pred. No. 9.6e-53;
  Matches  411;  Conservative    0;  Mismatches  235;  Indels    0;  Gaps    0;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156

Db         99 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 158

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216

Db        159 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 218

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276

Db        219 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 278

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336

Db        279 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 338

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396

Db        339 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 398

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456

Db        399 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 458

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516

Db        459 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 518

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576

Db        519 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 578

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636

Db        579 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 638

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTGCTGCTGATCCTGCTGTG 696

Db        639 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCTGATGAGACAGTT 698

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742

Db        699 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 744


RESULT 9
AF118924
LOCUS       AF118924                998 bp    mRNA    linear   PLN 02-JAN-2000
DEFINITION  Papaver somniferum glutathione S-transferase 1 mRNA, complete cds.
ACCESSION   AF118924
VERSION     AF118924.1  GI:6652869
KEYWORDS    .
SOURCE      Papaver somniferum (opium poppy)
  ORGANISM  Papaver somniferum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
            Papaveraceae; Papaver.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   Facchini,P.J. and Yu,M.
  TITLE     Molecular cloning and characterization of a glutathione
            S-transferase gene family from opium poppy
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 998)
  AUTHORS   Facchini,P.J. and Yu,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1999) Biological Sciences, University of Calgary,
            2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada
FEATURES             Location/Qualifiers
     source          1. .998
                     /organism="Papaver somniferum"
                     /mol_type="mRNA"
                     /cultivar="Marianne"
                     /db_xref="taxon:3469"
     CDS             78. .779
                     /EC_number="2.5.1.18"
                     /note="GST1"
                     /codon_start=1
```

/product="glutathione S-transferase 1"
/protein_id="AAF22517.1"
/db_xref="GI:6652870"
/translation="MAGSGSEEVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSK
SELLLKSNPIYKKMPVLIHGDKPICESMIIVQYIDDVWASAGHSIIPSDPYDASIARF
WATYIDDKFFPSLMGIAKSKDAEEKKAAIEQAIAAFGILEEAYQKTSKGKDFFGGEKI
GYVDIAFGCYVGWIRVTEKMNGIKLFDEEKVPGLTKWAEKFCADETVKSVMPETDALM
EFAKKIFGSKPPPSN"

ORIGIN

Query Match 29.4%; Score 270; DB 8; Length 998;
Best Local Similarity 63.6%; Pred. No. 9.6e-53;
Matches 411; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```
Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
Db         99 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 158

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
Db        159 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 218

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
Db        219 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 278

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
Db        279 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 338

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
Db        339 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 398

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
Db        399 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 458

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
Db        459 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 518

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
Db        519 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 578

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
Db        579 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 638

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
Db        639 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCTGATGAGACAGTT 698

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742
Db        699 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 744
```

RESULT 10
AX203247
LOCUS AX203247 947 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 3 from Patent WO0153501.
ACCESSION AX203247
VERSION AX203247.1 GI:15392618
KEYWORDS .
SOURCE Papaver somniferum (opium poppy)
ORGANISM Papaver somniferum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Papaveraceae; Papaver.
REFERENCE 1
AUTHORS Facchini,P.J.
TITLE Glutathione-s-transferase nucleic acids and polypeptides and
methods of use thereof
JOURNAL Patent: WO 0153501-A 3 26-JUL-2001;
Facchini, Peter James (CA)
FEATURES Location/Qualifiers
source 1. .947
/organism="Papaver somniferum"
/mol_type="unassigned DNA"
/db_xref="taxon:3469"

ORIGIN

Query Match 29.2%; Score 268.4; DB 6; Length 947;
Best Local Similarity 63.5%; Pred. No. 2.3e-52;
Matches 410; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

```
Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
Db         48 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 107

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
Db        108 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 167

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
Db        168 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 227

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
Db        228 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
Db        288 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
Db        348 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
Db        408 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
Db        468 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
Db        528 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
Db        588 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACGGTT 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742
Db        648 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 693
```

RESULT 11
AF118925
LOCUS AF118925 947 bp mRNA linear PLN 02-JAN-2000
DEFINITION Papaver somniferum glutathione S-transferase 2 mRNA, complete cds.
ACCESSION AF118925
VERSION AF118925.1 GI:6652871
KEYWORDS .
SOURCE Papaver somniferum (opium poppy)
ORGANISM Papaver somniferum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Papaveraceae; Papaver.
REFERENCE 1 (bases 1 to 947)

```
  AUTHORS   Facchini,P.J. and Yu,M.
  TITLE     Molecular cloning and characterization of a glutathione
            S-transferase gene family from opium poppy
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 947)
  AUTHORS   Facchini,P.J. and Yu,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1999) Biological Sciences, University of Calgary,
            2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada
FEATURES             Location/Qualifiers
     source          1. .947
                     /organism="Papaver somniferum"
                     /mol_type="mRNA"
                     /cultivar="Marianne"
                     /db_xref="taxon:3469"
     CDS             27. .728
                     /EC_number="2.5.1.18"
                     /note="GST2"
                     /codon_start=1
                     /product="glutathione S-transferase 2"
                     /protein_id="AAF22518.1"
                     /db_xref="GI:6652872"
                     /translation="MAGSGSEEVKILGGWPSPFVMRPRIALNIKSVKYYLLEETFGSK
                     SELLLKSNPIYKKMPVLIHGDKPICESMIIVQYIDDVWASAGHSIIPSDPYDASIARF
                     WATYIDDKFFPSLMGIAKSKDAEEKKAAIEQAIAAFGILEEAYQKTSKGKDFFGGEKI
                     GYVDIAFGCYVGWIRVTEKMNGIKLFDEEKVPGLTKWAEKFCADETVKSVMPETDALM
                     EFAKKIFGSKPPPSN"
ORIGIN

  Query Match           29.2%;  Score 268.4;  DB 8;  Length 947;
  Best Local Similarity   63.5%;  Pred. No. 2.3e-52;
  Matches  410;  Conservative    0;  Mismatches  236;  Indels    0;  Gaps    0;

Qy          97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
               ||  ||| | |||| |||| |||  ||||||||||    |||||      ||||| ||
Db          48 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 107

Qy         157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
               ||| | ||     |  | ||| |  || ||||||||||  ||   |  ||||||||| |
Db         108 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 167

Qy         217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
               ||||| || || || ||| |  ||||||| || || || ||  | || || ||||||  |
Db         168 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 227

Qy         277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
               || ||||||||    ||||| ||| ||||||| ||||| || ||| |       || ||
Db         228 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 287

Qy         337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
               |||||| | ||     || | ||||||    | | ||| || || ||||   | |||||
Db         288 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 347

Qy         397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
               ||||||||||   || |||| || |   |  ||  |  || | | ||||  || | |||
Db         348 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 407

Qy         457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
               ||   | | |||| || |||   | ||   |||  |   | |||||||  |   | ||
Db         408 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 467

Qy         517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
                  |||| |||| || |  | |||||| ||||| |  |||||  | ||| || ||||  |||
Db         468 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 527

Qy         577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
               ||  | | | |  | |  ||||||||  |||||||  |||||  | ||  |||  |  |||||
Db         528 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 587

Qy         637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
               |||   ||    || ||  | ||   |||||||||  |  |||  |||  |||    | ||
Db         588 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACGGTT 647

Qy         697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742
               ||     || |||| || |||||    ||  | |||||||| ||||
Db         648 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 693


RESULT 12
AY091102
LOCUS       AY091102                962 bp    mRNA    linear   PLN 18-SEP-2002
DEFINITION  Arabidopsis thaliana At1g10370 mRNA sequence.
ACCESSION   AY091102
VERSION     AY091102.1  GI:20268697
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 962)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Full Length cDNA Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 962)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
            Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.

            Annotation is based on the January 2002 version of the Arabidopsis
            genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1. .962
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="RAFL09-92-A11 (R19869)"
                     /ecotype="Columbia"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert."
     gene            1. .962
                     /gene="At1g10370"
```

```
     misc_difference 1..2
                     /gene="At1g10370"
                     /note="compared to genomic sequence"
                     /replace="ata"
     misc_feature    87..770
                     /gene="At1g10370"
                     /note="This is a potential full length cDNA corresponding
                     to At1g10370.  The cDNA contains a single base pair
                     insertion at position 564 resulting in a 3' elongated
                     ORF."
                     /evidence=experimental
     misc_difference 564
                     /gene="At1g10370"
                     /note="compared to genomic sequence"
     misc_difference 946
                     /gene="At1g10370"
                     /note="not present in genomic sequence"
ORIGIN

  Query Match             28.9%;  Score 265.4;  DB 8;  Length 962;
  Best Local Similarity   61.5%;  Pred. No. 1.1e-51;
  Matches  425;  Conservative    0;  Mismatches  266;  Indels    0;  Gaps    0;

Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
Db         87 ATGGCAAGCAGCGACGTGAAGCTGATCGGTGCATGGGCGAGTCCCTTTGTGATGAGGCCG 146

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
Db        147 AGGATTGCTCTAAACCTCAAGTCTGTCCCCTACGAGTTCCTCCAAGAGACGTTTGGGTCT 206

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
Db        207 AAGAGCGAGTTGCTTCTTAAATCAAACCCGGTTCACAAGAAGATCCCGGTTCTGCTTCAT 266

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
Db        267 GCTGATAAACCGGTGAGTGAGTCCAACATCATCGTTGAGTATATCGATGACACTTGGAGC 326

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
Db        327 TCATCTGGACCGTCCATTCTCCCTTCCGATCCTTACGATCGGGCCATGGCTCGGTTCTGG 386

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
Db        387 GCTGCTTACATCGACGAAAAGTGGTTTGTCGCTCTAAGAGGTTTCCTAAAAGCCGGAGGA 446

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
Db        447 GAAGAAGAGAAGAAAGCTGTGATAGCTCAACTAGAAGAAGGGAATGCGTTTCTGGAGAAG 506

Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
Db        507 GCGTTCATTGATTGCAGCAAAGGAAAACCGTTCTTCAACGGTGACAACATCGGTTACCTC 566

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
Db        567 GACATTGCTCTCGGGTGCTTCTTGGCTTGGTTGAGAGTCACCGAGTTAGCAGTCAGCTAT 626

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
Db        627 AAAATTCTTGATGAGGCCAAGACACCTTCTTTGTCCAAATGGGCTGAGAATTTCTGTAAT 686

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATT 744
Db        687 GATCCCGCTGTAAAACCTGTCATGCCCGAGACTGCAAAGCTTGCTGAATTCGCAAAGAAG 746

Qy        745 CTTCAGCTAAAATGGGCTGCTGCAGCTGCTG 775
Db        747 ATCTTTCCTAAGCCGCAGGCCTAAGTTGTTG 777

RESULT 13
AF159229
LOCUS       AF159229                 988 bp    mRNA    linear   PLN 03-FEB-2000
DEFINITION  Gossypium hirsutum glutathione S-transferase mRNA, complete cds.
ACCESSION   AF159229
VERSION     AF159229.1  GI:6856102
KEYWORDS    .
SOURCE      Gossypium hirsutum (upland cotton)
  ORGANISM  Gossypium hirsutum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE   1  (bases 1 to 988)
  AUTHORS   Kang,W.H., Yamamoto,E. and Allen,D.R.
  TITLE     Cloning and characterization of glutathione S-transferase (GST)
            gene from cotton plants (Gossypium hirsutum L. cv. Coker 312)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 988)
  AUTHORS   Kang,W.H., Yamamoto,E. and Allen,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-1999) Plant and Soil Science, Texas Tech
            University, 4th and Flint, Lubbock, TX 79409, USA
FEATURES             Location/Qualifiers
     source          1..988
                     /organism="Gossypium hirsutum"
                     /mol_type="mRNA"
                     /db_xref="taxon:3635"
                     /tissue_type="fiber"
     CDS             14..859
                     /note="GST"
                     /codon_start=1
                     /product="glutathione S-transferase"
                     /protein_id="AAF29773.1"
                     /db_xref="GI:6856103"
                     /translation="MQLETYWMLQMKAVTSVRKCEVSSLGFSFNLYINTCRVKTFVTL
                     KRNLTMAKSDAEVKVLGTWASPFVMRVRIALNIKSVAYEFLQERLWEGKSELLLKSNP
                     VHRKVPVLIHGDDTICESLIIVQYIDEVWPSVPILPSDPHERAAARFWAAYLDDKWFP
                     SLRAIGMAEGEDARKAAIGQVEEGLMLLEEAFGKCSQGQALFGKDQIGYLDITFGCFL
                     GWLRVTEKMSGIKLLNEINTPALLKWANRFCNDAAVKDVMPETEKLAEFAKMLRGRVR
                     ATPTS"
ORIGIN

  Query Match             28.8%;  Score 264.8;  DB 8;  Length 988;
  Best Local Similarity   64.6%;  Pred. No. 1.6e-51;
  Matches  430;  Conservative    0;  Mismatches  227;  Indels    9;  Gaps    2;

Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
Db        167 AAGAGTGATGCCGAAGTGAAAGTTTTGGGCACATGGGCGAGTCCATTCGTGATGAGGGTA 226

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTT---GAAT 201
Db        227 AGGATTGCCCTGAACATCAAGTCTGTGGCCTATGAATTTCTTCAGGAGAGGTTATGGGAA 286

Qy        202 CCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTC 261
Db        287 GGCAAAAGCGAGCTTCTCCTCAAATCAAACCCAGTTCACAGAAAGGTCCCTGTCCTCATT 346

Qy        262 CATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGG 321
Db        347 CATGGTGATGACACCATCTGTGAGTCTCTCATTATTGTCCAATACATCGATGAGGTTTGG 406

Qy        322 TCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTT 381
Db        407 CC------TTCTGTCCCCATTCTTCCTTCTGATCCCCATGAACGTGCCGCCGCTCGTTTT 460

Qy        382 TGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAA 441
Db        461 TGGGCTGCTTATCTTGACGACAAGTGGTTCCCTTCCTTGAGAGCTATTGGTATGGCTGAA 520

Qy        442 GATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAA 501
Db        521 GGAGAGGATGCAAGGAAAGCAGCCATAGGGCAAGTGGAGGAAGGGCTGATGCTGTTAGAG 580
```

```
Qy        502 GAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTT 561
              || |  ||   ||| |||||  ||||| |||| |  || ||    |||   |||||||
Db        581 GAGGCATTTGGCAAATGCAGCCAAGGGCAGGCATTGTTTGGGAAGGATCAAATTGGATAT 640

Qy        562 GTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAA 621
              |||||||    |||||  |||||||| | |    | |||||  ||||||| |||| ||
Db        641 CTTGACATAACATTTGGGTGCTTTTTGGGGTGGCTAAGAGTGACAGAGAAGATGAGTGGG 700

Qy        622 AGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCT 681
              |  ||  ||||| ||||||  ||    |||| | ||     |||||| | |  ||
Db        701 ATCAAGCTGCTTAATGAAATCAACACTCCTGCTCTGCTTAAATGGGCTAACAGATTCTGC 760

Qy        682 GCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAG 741
              |||| | || ||||| |   |  |||| ||||| ||||||||||  |||||||| |||
Db        761 AATGATGCCGCCGTGAAAGATGTCATGCCTGAGACCGAAAAGCTTGCGGAGTTTGCTAAG 820

Qy        742 ATTCTT 747
              || |||
Db        821 ATGCTT 826


RESULT 14
AF288191
LOCUS       AF288191                 684 bp    mRNA    linear   PLN 06-NOV-2000
DEFINITION  Arabidopsis thaliana chromosome I glutathione S-transferase (GST30)
            mRNA, complete cds.
ACCESSION   AF288191
VERSION     AF288191.1  GI:11096015
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 684)
  AUTHORS   Wagner,U. and Mauch,F.
  TITLE     Analysis of the glutathione S-transferase family in Arabidopsis
            thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 684)
  AUTHORS   Wagner,U. and Mauch,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2000) Department of Biology, University of
            Fribourg, 3, rte Albert Gockel, Fribourg CH- 1700, Switzerland
FEATURES             Location/Qualifiers
     source          1. .684
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="I"
                     /ecotype="Columbia"
     gene            1. .684
                     /gene="GST30"
     CDS             1. .684
                     /gene="GST30"
                     /codon_start=1
                     /product="glutathione S-transferase"
                     /protein_id="AAG30140.1"
                     /db_xref="GI:11096016"
                     /translation="MASSDVKLIGAWASPFVMRPRIALNLKSVPYEFLQETFGSKSEL
                     LLKSNPVHKKIPVLLHADKPVSESNIIVEYIDDTWSSSGPSILPSDPYDRAMARFWAA
                     YIDEKWFVALRGFLKAGGEEEKKAVIAQLEEGNAFLEKAFIDCSKGKPFFNGDNIGYL
                     DIALGCFLAWLRVTELAVSYKILDEAKTPSLSKWAENFCNDPAVKPVMPETAKLAEFA
                     KKIFPKPQA"
ORIGIN

  Query Match           28.8%;  Score 264.4;  DB 8;  Length 684;
  Best Local Similarity   62.6%;  Pred. No. 2e-51;
  Matches  412;  Conservative    0;  Mismatches  246;  Indels    0;  Gaps    0;

Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
              |||||    || ||| ||| |||  | ||||| |||   |||||  ||||   |||||  |
Db          1 ATGGCAAGCAGCGACGTGAAGCTGATCGGTGCATGGGCGAGTCCCTTTGTGATGAGGCCG 60
```

```
Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
              ||||||  || |||||||||  | |     || ||| |  |  |||||||  ||      |
Db         61 AGGATTGCTCTAAACCTCAAGTCTGTCCCCTACGAGTTCCTCCAAGAGACGTTTGGGTCT 120

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
              || || ||  |||||||||||| || ||||| || |||||||| ||||| ||| |  | |||
Db        121 AAGAGCGAGTTGCTTCTTAAATCAAACCCGGTTCACAAGAAGATCCCGGTTCTGCTTCAT 180

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
              |  ||||||    |  |||| ||    ||||| |||||||| || |||||   ||||  |
Db        181 GCTGATAAACCGGTGAGTGAGTCCAACATCATCGTTGAGTATATCGATGACACTTGGAGC 240

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
                ||  |  ||||| || ||      || | || ||||| || |  || || || |||
Db        241 TCATCTGGACCGTCCATTCTCCCTTCCGATCCTTACGATCGGGCCATGGCTCGGTTCTGG 300

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
              | | |||||||||| || |||||| ||    | | | | || | |||    | | ||
Db        301 GCTGCTTACATCGACGAAAAGTGGTTTGTCGCTCTAAGAGGTTTCCTAAAAGCCGGAGGA 360

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
              || || |  |||||        | |  |||   ||||||| |  || |    |||| |
Db        361 GAAGAAGAGAAGAAAGCTGTGATAGCTCAACTAGAAGAAGGGAATGCGTTTCTGGAGAAG 420

Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
              | |||||    | |||||  |||| ||  | |  |||   || || |  || || |   |
Db        421 GCGTTCATTGATTGCAGCAAAGGAAAACCGTTCTTCAACGGTGACAACATCGGTTACCTC 480

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
              |||||||  | |  ||  |||| |||   ||    | |||||| | |||
Db        481 GACATTGCTCTCGGGTGCTTCTTGGCTTGGTTGAGAGTCACCGAGTTAGCAGTCAGCTAT 540

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
              ||| | ||||||||  | |||   |||  |||| ||  |||||||| | |||   | |
Db        541 AAAATTCTTGATGAGGCCAAGACACCTTCTTTGTCCAAATGGGCTGAGAATTTCTGTAAT 600

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742
              |||||  |||||  ||      |  |||| |||||||  |||||||| |||  || |||||||
Db        601 GATCCCGCTGTAAAACCTGTCATGCCCGAGACTGCAAAGCTTGCTGAATTCGCAAAGA 658


RESULT 15
AB039930
LOCUS       AB039930                 964 bp    mRNA    linear   PLN 28-AUG-2001
DEFINITION  Arabidopsis thaliana ERD9 mRNA for glutathione S-transferase,
            complete cds.
ACCESSION   AB039930
VERSION     AB039930.1  GI:15375407
KEYWORDS    glutathione S-transferase.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Kiyosue,T., Yamaguchi-Shinozaki,K. and Shinozaki,K.
  TITLE     Cloning of cDNAs for genes that are early-responsive to dehydration
            stress (ERDs) in Arabidopsis thaliana L.: identification of three
            ERDs as HSP cognate genes
  JOURNAL   Plant Mol. Biol. 25 (5), 791-798 (1994)
  MEDLINE   94355652
   PUBMED   8075396
REFERENCE   2  (bases 1 to 964)
  AUTHORS   Shinozaki,K., Yamaguchi-shinozaki,K. and Takahashi,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) Kazuo Shinozaki, RIKEN, Laboratory of Plant
            Molecular Biology; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            (E-mail:sinozaki@rtc.riken.go.jp, Tel:+81-298-36-4359)
FEATURES             Location/Qualifiers
     source          1. .964
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
```

```
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="ERD9"
     gene            1. .964
                     /gene="ERD9"
     CDS             70. .753
                     /gene="ERD9"
                     /note="early-responsive to dehydration"
                     /codon_start=1
                     /product="glutathione S-transferase"
                     /protein_id="BAB63917.1"
                     /db_xref="GI:15375408"
                     /translation="MASSDVKLIGAWASPFVMRPRIALNLKSVPYEFLQETFGSKSEL
                     LLKSNPVHKKIPVLLHADKPVSESNIIVEYIDDTWSSSGPSILPSDPYDRAMARFWAA
                     YIDEKWFVALRGFLKAGGEEEKKAVIAQLEEGNAFLEKAFIDCSKGKPSFNGDNIGYL
                     NIALGCFLAWLRVTELAVSYKILDEAKTPSLSKWAENFCNDPAVKPVMPETAKLAEFA
                     KKIFPKPQA"
ORIGIN

  Query Match           28.7%;  Score 263.8;  DB 8;  Length 964;
  Best Local Similarity 61.4%;  Pred. No. 2.7e-51;
  Matches  424;  Conservative    0;  Mismatches  267;  Indels    0;  Gaps    0;

Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
              |||||     || ||| ||| |||   | ||||| |||    ||||| ||||   |||||  |
Db         70 ATGGCAAGCAGCGACGTGAAGCTGATCGGTGCATGGGCGAGTCCCTTTGTGATGAGGCCG 129

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
                |||||| || ||||||||| | |    || ||| |  |  |||||| ||     |
Db        130 AGGATTGCTCTAAACCTCAAGTCTGTCCCCTACGAGTTCCTCCAAGAGACGTTTGGGTCT 189

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
              || || || |||||||||||| || ||||| || |||||||| ||||| ||| |  | |||
Db        190 AAGAGCGAGTTGCTTCTTAAATCAAACCCGGTTCACAAGAAGATCCCGGTTCTGCTTCAT 249

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
              |  ||||||     |  |||| ||    ||||| |||||||| || |||||    ||||  |
Db        250 GCTGATAAACCGGTGAGTGAGTCCAACATCATCGTTGAGTATATCGATGACACTTGGAGC 309

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
                ||  |  ||||| || ||     || | || ||||| || |  || || || |||
Db        310 TCATCTGGACCGTCCATTCTCCCTTCCGATCCTTACGATCGGGCCATGGCTCGGTTCTGG 369

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
              | | |||||||||| || |||||| ||     |  | | | || | |||     | | ||
Db        370 GCTGCTTACATCGACGAAAAGTGGTTTGTCGCTCTAAGAGGTTTCCTAAAAGCCGGAGGA 429

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
              || || |  |||||        | |  |||   ||||||| |  || |    ||||  |
Db        430 GAAGAAGAGAAGAAAGCTGTGATAGCTCAACTAGAAGAAGGGAATGCGTTTCTGGAGAAG 489

Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
              | |||||    | |||||  |||| ||  | |  |||   || || |  || || |   |
Db        490 GCGTTCATTGATTGCAGCAAAGGAAAACCGTCCTTCAACGGTGACAACATCGGTTACCTC 549

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
              ||||||  | | ||  |||| |||   ||    | |||||| | |||
Db        550 AACATTGCTCTCGGGTGCTTCTTGGCTTGGTTGAGAGTCACCGAGTTAGCAGTCAGCTAT 609

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
              ||| | |||||||||  | |||   |||  |||| ||  ||||||||| | |||   |  |
Db        610 AAAATTCTTGATGAGGCCAAGACACCTTCTTTGTCCAAATGGGCTGAGAATTTCTGTAAT 669

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATT 744
              ||||| ||||| ||      |  |||| |||||||| |||||||| ||| || |||||||
Db        670 GATCCCGCTGTAAAACCTGTCATGCCCGAGACTGCAAAGCTTGCTGAATTCGCAAAGAAG 729

Qy        745 CTTCAGCTAAAATGGGCTGCTGCAGCTGCTG 775
              |   |  ||   |  ||   || || ||
Db        730 ATCTTTCCTAAGCCGCAGGCCTAAGTTGTTG 760

Search completed: October 13, 2004, 05:00:43
```

Job time : 3931 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 12, 2004, 18:31:26 ; Search time 448 Seconds
(without alignments)
8705.013 Million cell updates/sec

Title: US-10-088-945A-14
Perfect score: 918
Sequence: 1 cataaaactccacatttcct.....aatatagtaaaaaaaaaaaa 918

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	918	100.0	918	4	AAF57513	Aaf57513 Glycine m
2	620.2	67.6	895	3	AAA53396	Aaa53396 Clone se3
3	620.2	67.6	895	3	AAZ94952	Aaz94952 Soybean g
4	620.2	67.6	895	3	AAA59460	Aaa59460 Nucleotid
5	620.2	67.6	895	4	AAF31575	Aaf31575 Soybean t
6	620.2	67.6	895	4	AAF57526	Aaf57526 Glycine m
7	288	31.4	896	4	AAC86194	Aac86194 GST3 nucl
8	270	29.4	998	4	AAC86192	Aac86192 GST1 nucl
9	268.4	29.2	947	4	AAC86193	Aac86193 GST2 nucl
10	240	26.1	684	6	ABZ14871	Abz14871 Arabidops
11	232.2	25.3	705	6	ABZ14853	Abz14853 Arabidops
c 12	228.2	24.9	817	6	ABN98975	Abn98975 Arabidops
13	226.4	24.7	955	3	AAC41255	Aac41255 Arabidops
14	202.2	22.0	863	3	AAC51154	Aac51154 Arabidops
15	197.4	21.5	866	3	AAC35433	Aac35433 Arabidops
16	193.6	21.1	513	6	ABZ13860	Abz13860 Arabidops
17	184	20.0	810	7	ADA70384	Ada70384 Rice gene
18	181.8	19.8	684	6	ABZ12890	Abz12890 Arabidops
19	181.8	19.8	881	3	AAC47459	Aac47459 Arabidops
c 20	181.4	19.8	824	6	ABN98949	Abn98949 Arabidops
21	181.2	19.7	702	7	ADA70491	Ada70491 Rice gene
22	178.4	19.4	696	7	ADA69491	Ada69491 Rice gene
23	173.6	18.9	708	7	ADA70490	Ada70490 Rice gene
24	172.2	18.8	881	3	AAC34181	Aac34181 Arabidops
25	171	18.6	834	7	ADA69909	Ada69909 Rice gene
26	167.4	18.2	970	3	AAA76274	Aaa76274 Maize glu
27	166.2	18.1	895	2	AAZ07557	Aaz07557 Maize glu
28	166.2	18.1	895	3	AAZ94797	Aaz94797 Maize cla
29	166.2	18.1	895	3	AAA76262	Aaa76262 Maize glu
30	164	17.9	1068	3	AAA76287	Aaa76287 Maize glu
31	160.4	17.5	702	7	ADA70385	Ada70385 Rice gene
32	157.6	17.2	904	3	AAA76285	Aaa76285 Maize glu
33	156.8	17.1	1143	3	AAC38593	Aac38593 Arabidops
34	156	17.0	960	3	AAA76279	Aaa76279 Maize glu
35	152.4	16.6	934	3	AAA76278	Aaa76278 Maize glu
36	152	16.6	708	7	ADA69765	Ada69765 Rice gene
37	142.6	15.5	900	3	AAA76281	Aaa76281 Maize glu
38	142.6	15.5	926	2	AAX25151	Aax25151 Wheat Typ
c 39	142.2	15.5	736	8	ACL25466	Acl25466 DNA clone
40	140	15.3	841	3	AAA59485	Aaa59485 Nucleotid
41	139.8	15.2	860	3	AAA76271	Aaa76271 Maize glu
42	138.4	15.1	840	4	AAF57512	Aaf57512 Glycine m
c 43	133.8	14.6	653	8	ACL25409	Acl25409 DNA clone
c 44	132.2	14.4	642	8	ACL25401	Acl25401 DNA clone
c 45	131.8	14.4	754	8	ACL25435	Acl25435 DNA clone

ALIGNMENTS

RESULT 1
AAF57513
ID AAF57513 standard; DNA; 918 BP.
XX
AC AAF57513;
XX
DT 11-JUN-2001 (first entry)
XX
DE Glycine max glutathione-S-transferase (GST) 3.3 encoding DNA.
XX
KW Soybean; glutathione-S-transferase; GST; homoglutathione synthetase; HGS;
KW beta-alanine; gamma-glutamylcysteine; herbicide; ds.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 85. .783
FT /*tag= a
FT /product= "GST 3.3"
XX
PN WO200121770-A2.
XX
PD 29-MAR-2001.
XX
PF 18-SEP-2000; 2000WO-GB003573.
XX
PR 21-SEP-1999; 99GB-00022346.
XX
PA (ZENE) ZENECA LTD.
XX
PI Andrews CJ, Jepson I, Townson JK, Edwards R, Cummins I;
PI Skipsey M;
XX
DR WPI; 2001-257978/26.
DR P-PSDB; AAB62226.
XX
PT Novel glutathione-S-transferase and homoglutathione synthetase sequences
PT from soybean for producing plants which are resistant and tolerant to
PT herbicide comprising fomesafen and/or acifluorfen.
XX
PS Claim 4; Page 47; 64pp; English.
XX
CC The invention relates to new soybean glutathione-S-transferase (GST) and
CC a homoglutathione synthetase (HGS). The HGS is capable of catalyzing the
CC addition of beta-alanine onto gamma-glutamylcysteine. Polynucleotides
CC encoding the proteins of the invention are useful for producing plants

```
CC   which are resistant and/or tolerant to a herbicide comprising fomesafen
CC   and/or acifluorfen. Methods of the invention are useful for providing
CC   plants with further desired agronomic trait, especially resistant to a
CC   herbicide, comprising glyphosate or its salt. Further desired agronomic
CC   traits include insect resistance, nematode resistance, stress tolerance,
CC   altered field, altered nutritional value, altered quality or any other
CC   desirable agronomic trait. GST or its variant is also useful as a
CC   selectable marker gene. The present sequence represents a Glycine max GST
CC   3.3 (also referred to as GST 3.6) protein encoding DNA
XX
SQ   Sequence 918 BP; 281 A; 151 C; 212 G; 274 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 918;  DB 4;  Length 918;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-222;
  Matches  918;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATAAAACTCCACATTTCCTGCTGAGTAACCTAACAAAACAAACACAATATTGCTCCGTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATAAAACTCCACATTTCCTGCTGAGTAACCTAACAAAACAAACACAATATTGCTCCGTG 60

Qy         61 TTTGACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTGACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGG 120

Qy        121 TTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 180

Qy        181 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240

Qy        241 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 300

Qy        301 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 360

Qy        361 GATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 420

Qy        421 AAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAA 480

Qy        481 GAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTC 540

Qy        541 GGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGA 600

Qy        601 GTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACC 660

Qy        661 CTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAA 720

Qy        721 AAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAG 780

Qy        781 TAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATA 840

Qy        841 TCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGATAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGATAA 900

Qy        901 TATAGTAAAAAAAAAAAA 918
              ||||||||||||||||||
Db        901 TATAGTAAAAAAAAAAAA 918


RESULT 2
AAA53396
ID   AAA53396 standard; cDNA; 895 BP.
XX
AC   AAA53396;
XX
DT   04-OCT-2000  (first entry)
XX
DE   Clone se3.03b09 nucleotide sequence encoding GST type III.
XX
KW   Soybean; glutathione-S-transferase; GST; detoxify; herbicide; stress;
KW   transgenic plant; tolerant; plant breeding; ss.
XX
OS   Glycine max.
XX
PN   US6063570-A.
XX
PD   16-MAY-2000.
XX
PF   05-SEP-1997;   97US-00924747.
XX
PR   05-SEP-1997;   97US-00924747.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Mcgonigle B,  O'keefe DP;
XX
DR   WPI; 2000-375487/32.
DR   P-PSDB; AAB03734.
XX
PT   New Glutathione-S-Transferase enzymes and isolated nucleic acid fragments
PT   encoding them, useful for detoxifying xenobiotic compounds in plants and
PT   seeds, as well as in producing transgenic plants that are herbicide-
PT   resistant.
XX
PS   Claim 2; Col 33-36; 36pp; English.
XX
CC   This sequence was isolated from a soybean clone, and encodes a
CC   Glutathione-S-Transferase (GST). The invention relates to isolated
CC   nucleic acid fragments (see AAA53393-A53406) which encode soybean GST
CC   polypeptides (AAB03731-B03744). GSTs are a family of enzymes which
CC   catalyse the conjugation of glutathione, homoglutathione and other
CC   glutathione-like analogues, to a large range of hydrophobic,
CC   electrophilic compounds. GSTs have been implicated in the detoxification
CC   of certain herbicides. The GST nucleotide sequences are useful in the
CC   construction of herbicide-tolerant transgenic plants, plants that are
CC   tolerant to a wide variety of stresses, or plants in which the GST
CC   enzymes are present at higher or lower levels than they are normally. The
CC   nucleic acid fragments are also useful as probes for genetically and
CC   physically mapping the genes that they are part of, and as markers for
CC   traits linked to expression of the enzymes. This will be useful in plant
CC   breeding in order to develop lines with desired phenotypes or in the
CC   identification of mutants. The soybean GST enzymes are used to detoxify
CC   xenobiotic compounds in plants and seeds. The enzymes are also useful as
CC   targets to facilitate design and/or identify inhibitors of the enzymes
CC   that may be used as herbicides or herbicide synergists. The GST enzymes
CC   produced in the host cells, particularly in microbial host cells, are
CC   useful in preparing antibodies to the enzymes. These antibodies are
CC   useful for detecting the enzymes in situ in cells or in vitro in cell
CC   extracts
XX
SQ   Sequence 895 BP; 260 A; 146 C; 215 G; 274 T; 0 U; 0 Other;

  Query Match           67.6%;  Score 620.2;  DB 3;  Length 895;
  Best Local Similarity  88.0%;  Pred. No. 5.4e-147;
```

```
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| |||||||||  |  |||||||| |||||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || |||||||||||||||||||| ||||||||||||||||||  ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||| ||||||| ||||    |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||| |||||||| |||||||| |||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| |||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | |||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| ||||||| || ||||||  |||| | |||| || ||||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || |||||||||||||||||||||||||||||  || | |
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  ||||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |     | |  || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 3
AAZ94952
ID   AAZ94952 standard; cDNA; 895 BP.
XX
AC   AAZ94952;
XX
DT   01-AUG-2000  (first entry)
XX
DE   Soybean glutathione-S-transferase cDNA clone se3.03B09.
XX
KW   Soybean; glutathione-S-transferase; GST; xenobiotic; detoxification;
KW   transgenic plant; herbicide tolerance; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             36. .740
FT                   /*tag=  a
FT                   /EC_number= "2.5.1.18"
XX
PN   WO200018936-A1.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1998;   98WO-US020501.
XX
PR   30-SEP-1998;   98WO-US020501.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Mcgonigle B,  O'keefe DP;
XX
DR   WPI; 2000-317517/27.
DR   P-PSDB; AAY79515.
XX
PT   Nucleic acids encoding soybean glutathione-S-transferase enzymes useful
PT   for conferring herbicide resistance to plants.
XX
PS   Claim 2; Page 45; 76pp; English.
XX
CC   The present sequence is that of the cDNA insert in clone se3.03B09
CC   encoding soybean class III glutathione-S-transferase (GST, see AAY79515).
CC   The clone was isolated from a cDNA library prepared from soybean embryo.
CC   The invention provides soybean GST enzymes (see AAY79512-25) involved in
CC   the detoxification of xenobiotic compounds, especially herbicides, in
CC   plants and seeds. Chimeric genes encoding all or a portion of soybean GST
CC   enzymes, host cells, and methods of recombinant production of soybean GST
CC   enzymes are provided. The sequences are useful in the construction of
CC   herbicide-tolerant transgenic plants, in the recombinant production of
CC   GST enzymes, in the development of screening assays to identify compounds
CC   inhibitory to the GST enzymes (useful as herbicides or herbicide
CC   synergists), and in screening assays to identify chemical substrates of
CC   the GSTs
XX
SQ   Sequence 895 BP; 260 A; 146 C; 215 G; 274 T; 0 U; 0 Other;

  Query Match             67.6%;  Score 620.2;  DB 3;  Length 895;
  Best Local Similarity   88.0%;  Pred. No. 5.4e-147;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| |||||||||  |  |||||||| |||||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || |||||||||||||||||||| ||||||||||||||||||  ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||| ||||||| ||||    |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||| |||||||| |||||||| |||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| |||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467
```

```
Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | ||||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||   || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||||| |||||||||||||||||| ||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || ||||||||||||||||||||||||||||  || | ||
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  |||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |      | |   | || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868
```

RESULT 4
AAA59460
ID AAA59460 standard; DNA; 895 BP.
XX
AC AAA59460;
XX
DT 14-NOV-2000 (first entry)
XX
DE Nucleotide sequence of a soybean type III glutathione-S-transferase.
XX
KW Soybean; glutathione-S-transferase; GST; detoxification;
KW xenobiotic compound; herbicide-tolerance; transgenic plant;
KW herbicide synergist; ss.
XX
OS Glycine max.
XX
FH Key   Location/Qualifiers
FT CDS   36. .740
FT   /*tag= a
FT   /product= "glutathione-S-transferase"
XX
PN WO200047728-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US003347.
XX
PR 10-FEB-1999; 99US-00247373.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Mcgonigle B, O'keefe DP;
XX
DR WPI; 2000-549144/50.
DR P-PSDB; AAB07829.
XX
PT Soybean glutathione-S-transferase polypeptides and polynucleotides used
PT to produce herbicide tolerant transgenic plants and to screen for
PT inhibitors or substrates of the enzyme.
XX
PS Claim 2; Page 58; 84pp; English.
XX
CC The present sequence encodes a soybean glutathione-S-transferase (GST)
CC enzyme. The enzyme is involved in the detoxification of xenobiotic
CC compounds in plants and seeds. The GST polynucleotides and polypeptides
CC are used for the production of herbicide-tolerant transgenic plants, and
CC for the development of screening assays to identify GST inhibitors and
CC substrates, which can be used as herbicide synergists. GST Gene specific
CC probes can be used in gene identification methods. The recombinant GST
CC enzymes can be used to produce enzyme specific antibodies which are used
CC to detect the enzymes in situ in cells or in vitro in cell extracts
XX
SQ Sequence 895 BP; 260 A; 146 C; 215 G; 274 T; 0 U; 0 Other;

```
  Query Match             67.6%;  Score 620.2;  DB 3;  Length 895;
  Best Local Similarity   88.0%;  Pred. No. 5.4e-147;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| ||||||||| |  |||||||| |||||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || ||||||||||||||||||| ||||||||||||||||||  ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||||| |||||||| |||||||| ||||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | ||||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||   || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||||| |||||||||||||||||| ||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || ||||||||||||||||||||||||||||  || | ||
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  |||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |      | |   | || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868
```

RESULT 5

```
AAF31575
ID   AAF31575 standard; DNA; 895 BP.
XX
AC   AAF31575;
XX
DT   09-APR-2001  (first entry)
XX
DE   Soybean type III GST cDNA #2.
XX
KW   Soybean; glutathione-S-transferase; herbicide; GST; ds.
XX
OS   Glycine max.
XX
PN   US6171839-B1.
XX
PD   09-JAN-2001.
XX
PF   22-APR-1999;   99US-00296715.
XX
PR   05-SEP-1997;   97US-00924747.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Mcgonigle B,  O'keefe DP;
XX
DR   WPI; 2001-136874/14.
XX
PT   Novel soybean glutathione-S-transferase enzymes useful as targets to
PT   facilitate design and/or identification of inhibitors of the enzyme, that
PT   are used as herbicides or herbicide synergists.
XX
PS   Claim 2; Col 33-36; 37pp; English.
XX
CC   The present invention relates to soybean glutathione-S-transferase
CC   proteins. The novel sequences are useful in the construction of herbicide
CC   tolerant transgenic plants, in the recombinant production of glutathione-
CC   S-transferase (GST) enzymes, in the development of screening assays to
CC   identify compounds inhibitory to the GST enzymes, and in screening assays
CC   to identify chemical substrates of the GSTs
XX
SQ   Sequence 895 BP; 260 A; 146 C; 215 G; 274 T; 0 U; 0 Other;

  Query Match             67.6%;  Score 620.2;  DB 4;  Length 895;
  Best Local Similarity   88.0%;  Pred. No. 5.4e-147;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy          97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
               ||||||| |||||||||  |  |||||||| ||||||||||||||||||||||||||||||
Db          51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy         157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
               ||||||||||||  || ||||||||||||||||||| ||||||||||||||||||  ||
Db         111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy         217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy         277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
               ||  |||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db         231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy         337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
               |||||||||||||||||| |||||||| |||||||| |||||||| || | |||||
Db         288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy         397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
               ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db         348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy         457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
               |||| |||||||||||||| |||||||| |||||||| | ||||||||||||||||||||
Db         408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy         517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
               | |||||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db         468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy         577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
               || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||||||||
Db         528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy         637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
               |||| |||| |||||||||||||| |||||||||||| ||||||||||||||||||||||
Db         588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy         697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
               |||||| |||| |||||||||||| |||||||||||||||||| ||||||||||||||||
Db         648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy         757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
               ||| ||||||||      || ||||||||||||||||||||||||||||||  || | |
Db         708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy         811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
               |||  |||||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db         767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy         871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
               |||||||||| |   ||||| | |      | |   | || |||
Db         826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 6
AAF57526
ID   AAF57526 standard; cDNA; 895 BP.
XX
AC   AAF57526;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Glycine max clone SE3.03B09 cDNA sequence.
XX
KW   Soybean; glutathione-S-transferase; GST; homoglutathione synthetase; HGS;
KW   beta-alanine; gamma-glutamylcysteine; herbicide; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             36. .740
FT                   /*tag=  a
XX
PN   WO200121770-A2.
XX
PD   29-MAR-2001.
XX
PF   18-SEP-2000; 2000WO-GB003573.
XX
PR   21-SEP-1999;   99GB-00022346.
XX
PA   (ZENE ) ZENECA LTD.
XX
PI   Andrews CJ,  Jepson I,  Townson JK,  Edwards R,  Cummins I;
PI   Skipsey M;
XX
DR   WPI; 2001-257978/26.
DR   P-PSDB; AAB62227.
XX
PT   Novel glutathione-S-transferase and homoglutathione synthetase sequences
PT   from soybean for producing plants which are resistant and tolerant to
PT   herbicide comprising fomesafen and/or acifluorfen.
XX
PS   Disclosure; Page 60-61; 64pp; English.
XX
CC   The invention relates to new soybean glutathione-S-transferase (GST) and
CC   a homoglutathione synthetase (HGS). The HGS is capable of catalyzing the
```

```
CC   addition of beta-alanine onto gamma-glutamylcysteine. Polynucleotides
CC   encoding the proteins of the invention are useful for producing plants
CC   which are resistant and/or tolerant to a herbicide comprising fomesafen
CC   and/or acifluorfen. Methods of the invention are useful for providing
CC   plants with further desired agronomic trait, especially resistant to a
CC   herbicide, comprising glyphosate or its salt. Further desired agronomic
CC   traits include insect resistance, nematode resistance, stress tolerance,
CC   altered field, altered nutritional value, altered quality or any other
CC   desirable agronomic trait. GST or its variant is also useful as a
CC   selectable marker gene. The present sequence represents a Glycine max
CC   clone SE3.03B09 cDNA sequence. The GST 3.3 cDNA sequence of the invention
CC   is specifically claimed to be not a part of the present sequence
XX
SQ   Sequence 895 BP; 260 A; 146 C; 215 G; 274 T; 0 U; 0 Other;

  Query Match           67.6%;  Score 620.2;  DB 4;  Length 895;
  Best Local Similarity  88.0%;  Pred. No. 5.4e-147;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| ||||||||| |  |||||||| |||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || |||||||||||||||||| ||||||||||||||||  ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||| ||||||| ||||    |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||| |||||||| |||||||| ||||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | |||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||| |||||||||||||||| ||||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || |||||||||||||||||||||||||||||  || | |
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  |||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |     | |   | || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 7
AAC86194
ID   AAC86194 standard; cDNA; 896 BP.
XX
AC   AAC86194;
XX
DT   19-SEP-2001  (first entry)
XX
DE   GST3 nucleic acid.
XX
KW   Glutathione-S-transferase; GST; opium poppy; conjugation;
KW   electrophilic compound; glutathione; GSH; detoxification; xenobiotic;
KW   transgenic plant; stress resistance; pathogen; herbicide; grazing pest;
KW   ss.
XX
OS   Papaver somniferum.
XX
FH   Key             Location/Qualifiers
FT   CDS             61. .762
FT                   /*tag=  a
FT                   /product= "GST3"
XX
PN   WO200153501-A2.
XX
PD   26-JUL-2001.
XX
PF   18-JAN-2001; 2001WO-IB000205.
XX
PR   18-JAN-2000; 2000US-0176708P.
XX
PA   (FACC/) FACCHINI P J.
XX
PI   Facchini PJ;
XX
DR   WPI; 2001-457612/49.
DR   P-PSDB; AAB47340.
XX
PT   Glutathione-s-transferase polypeptide(s) useful for producing transgenic
PT   plants comprising resistance to biotic and abiotic stress.
XX
PS   Claim 2; Page 10; 77pp; English.
XX
CC   This sequence encodes a glutathione-S-transferase (GST) polypeptide,
CC   GST3, from the opium poppy. GST's catalyse the conjugation of
CC   electrophilic compounds to glutathione (GSH). They have been shown to be
CC   involved in detoxification of xenobiotics. The novel GST's of the
CC   invention GST1, GST2 and GST3, show extensive homology with tau type
CC   GST's from a variety of plant species, especially within the N-terminal
CC   region. This domain has been shown to be responsible for recognition of
CC   GSH and the ability of the proteins to form dimers. The C-terminal domain
CC   is responsible for substrate specificity and varies greatly between other
CC   GST's. The GST's of the invention exhibit strong glutathione conjugating
CC   activity towards a model substrate, 1-chloro-2,4-dinitrobenzene that
CC   could be inhibited in the presence of hydroxycinnamic acid amines of
CC   tyramine. The expression of the GST polypeptide in a transgenic plant is
CC   useful to produce a plant which has increased stress resistance. The
CC   plant has increased resistance to pathogens, herbicides or grazing pests.
CC   The GST polypeptide is useful to identify binding agents inhibitors and
CC   substrates
XX
SQ   Sequence 896 BP; 304 A; 129 C; 192 G; 271 T; 0 U; 0 Other;

  Query Match           31.4%;  Score 288;  DB 4;  Length 896;
  Best Local Similarity  59.6%;  Pred. No. 8e-63;
  Matches  486;  Conservative    0;  Mismatches  330;  Indels    0;  Gaps    0;

Qy         96 GGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCT 155
              |||  | | | |||| |||| |||  ||||||||||   |||||      ||||| ||
Db         81 GGAGGTAAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACT 140
```

```
Qy        156 TAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATT 215
              ||| | ||     |  | ||| |  || ||||||||||| ||   |  |||||||||| |
Db        141 CAACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACT 200

Qy        216 GCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGT 275
              ||||| || || || ||| |  |||||||||| || ||| |  | ||||| ||||||
Db        201 TCTTCTGAAATCAAATCCTATTTACAAGAAAATTCCTGTTATGATTCATGGTGATAAACC 260

Qy        276 CATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCT 335
              ||| ||||||||     ||||| ||| ||||||| ||||| ||||||  |       ||  |
Db        261 CATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTTTGGGCTTCTGCTGGACA 320

Qy        336 CTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACAT 395
              ||  ||| | ||     || | ||||||     | | ||| || || ||||    | |||||
Db        321 TTCTATCATCCCGTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACAT 380

Qy        396 CGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAA 455
              ||||||||||    || |||| || |   |  ||  |  || | | ||||  || | |||
Db        381 TGATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAA 440

Qy        456 GAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAA 515
              ||   |  | |||| || |||   | ||   |||  |   |||||||||  |   | ||
Db        441 AAAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTGGAAGAAGCTTATCAGAA 500

Qy        516 GTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTT 575
              |||  |||| || | || ||||||  | |||  |  |||||||    |||| ||||  ||
Db        501 AACTAGTAAAGGAAAAGACTTTTTCGGGGAAGAAAAAATTGGATACATTGATATTGCATT 560

Qy        576 TGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGA 635
              |||  | | | |  |||  ||||||||  ||||||| ||||||| ||  |||  |  ||||
Db        561 TGGGTGTTATATAGGTTGGATTAGAGTTACAGAGAAAATGAATGGAATCAAACTATTTGA 620

Qy        636 TGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTGCTGCTGATCCTGCTGT 695
              ||||||  ||     ||  ||  | ||   |||||||||  |  |||  |||  |||    | ||
Db        621 TGAAACAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACAGT 680

Qy        696 GAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAA 755
              ||     || |||| || |||||    ||  | |||||||| ||||   |      |
Db        681 TAAATCTGTTATGCCTGAAACTGATGCTCTCATGGAGTTTGCTAAGAAGATCTTTGGATC 740

Qy        756 ATGGGCTGCTGCAGCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAATTTCA 815
                | || || |  |        ||  |  ||  | |||   |||  ||    ||  || |
Db        741 TAAGCCTCCTCCTTCAAACTAGAAAAAGTTGTTAACAATGAAATATCTTAGAGATGTTTA 800

Qy        816 AAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATATATTTA 875
              |   |||| |     ||        | ||  || ||| ||          ||| || | |
Db        801 AGCTTTGTGTTTGTTTTTTCAGTGTTGTGTAGCAATGCTTAAGAACTGTTTGTAGAAATG 860

Qy        876 AAAGTCAAAATAAATGTTATGATAATATAGTAAAAA 911
              |    || |  |    ||  | ||| | |  |||||
Db        861 ATCAAGAACAGTAGCTGTAAAAAAAAAAAAAAAAA 896


RESULT 8
AAC86192
ID   AAC86192 standard; cDNA; 998 BP.
XX
AC   AAC86192;
XX
DT   19-SEP-2001  (first entry)
XX
DE   GST1 nucleic acid.
XX
KW   Glutathione-S-transferase; GST; opium poppy; conjugation;
KW   electrophilic compound; glutathione; GSH; detoxification; xenobiotic;
KW   transgenic plant; stress resistance; pathogen; herbicide; grazing pest;
KW   ss.
XX
OS   Papaver somniferum.
XX
PN   WO200153501-A2.
XX
PD   26-JUL-2001.
XX
PF   18-JAN-2001; 2001WO-IB000205.
XX
PR   18-JAN-2000; 2000US-0176708P.
XX
PA   (FACC/) FACCHINI P J.
XX
PI   Facchini PJ;
XX
DR   WPI; 2001-457612/49.
XX
PT   Glutathione-s-transferase polypeptide(s) useful for producing transgenic
PT   plants comprising resistance to biotic and abiotic stress.
XX
PS   Claim 2; Page 8; 77pp; English.
XX
CC   This sequence encodes a glutathione-S-transferase (GST) polypeptide, GST1
CC   from the opium poppy. GST's catalyse the conjugation of electrophilic
CC   compounds to glutathione (GSH). They have been shown to be involved in
CC   detoxification of xenobiotics. The novel GST's of the invention GST1,
CC   GST2 and GST3, show extensive homology with tau type GST's from a variety
CC   of plant species, especially within the N-terminal region. This domain
CC   has been shown to be responsible for recognition of GSH and the ability
CC   of the proteins to form dimers. The C-terminal domain is responsible for
CC   substrate specificity and varies greatly between other GST's. The GST's
CC   of the invention exhibit strong glutathione conjugating activity towards
CC   a model substrate, 1-chloro-2,4-dinitrobenzene that could be inhibited in
CC   the presence of hydroxycinnamic acid amines of tyramine. The expression
CC   of the GST polypeptide in a transgenic plant is useful to produce a plant
CC   which has increased stress resistance. The plant has increased resistance
CC   to pathogens, herbicides or grazing pests. The GST polypeptide is useful
CC   to identify binding agents inhibitors and substrates
XX
SQ   Sequence 998 BP; 336 A; 143 C; 213 G; 306 T; 0 U; 0 Other;

  Query Match           29.4%;  Score 270;  DB 4;  Length 998;
  Best Local Similarity  63.6%;  Pred. No. 3e-58;
  Matches  411;  Conservative    0;  Mismatches  235;  Indels    0;  Gaps    0;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||  ||| | |||| ||||  |||   |||||||||   |||||       ||||| ||
Db         99 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 158

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              ||| | ||     |  | ||| |  || ||||||||||| ||    |  |||||||||| |
Db        159 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 218

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||| || || || ||| |  ||||||| || || || ||  | || || ||||||  |
Db        219 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 278

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || ||||||||     ||||| ||| ||||||| ||||| || ||| |       || ||
Db        279 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 338

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              ||||||  | ||     || | ||||||     | | ||| || || ||||    | |||||
Db        339 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 398

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||||||||    || |||| || |   |  ||  |  || | | ||||  || | |||
Db        399 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 458

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              ||   |  | |||| || |||   | ||   |||  |   | |||||||  |   | ||
Db        459 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 518

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
                 ||| |||| || |  | ||||||  |||||  |  |||||  | |||  || ||||  |||
Db        519 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 578

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              ||  | | | |  | |  ||||||||  |||||||  |||||  | ||  |||  |  |||||
```

```
Db        579 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 638

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696

Db        639 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCTGATGAGACAGTT 698

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742

Db        699 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 744


RESULT 9
AAC86193
ID   AAC86193 standard; cDNA; 947 BP.
XX
AC   AAC86193;
XX
DT   19-SEP-2001  (first entry)
XX
DE   GST2 nucleic acid.
XX
KW   Glutathione-S-transferase; GST; opium poppy; conjugation;
KW   electrophilic compound; glutathione; GSH; detoxification; xenobiotic;
KW   transgenic plant; stress resistance; pathogen; herbicide; grazing pest;
KW   ss.
XX
OS   Papaver somniferum.
XX
FH   Key             Location/Qualifiers
FT   CDS             27. .728
FT                   /*tag=  a
FT                   /product= "GST2"
XX
PN   WO200153501-A2.
XX
PD   26-JUL-2001.
XX
PF   18-JAN-2001; 2001WO-IB000205.
XX
PR   18-JAN-2000; 2000US-0176708P.
XX
PA   (FACC/) FACCHINI P J.
XX
PI   Facchini PJ;
XX
DR   WPI; 2001-457612/49.
DR   P-PSDB; AAB47339.
XX
PT   Glutathione-s-transferase polypeptide(s) useful for producing transgenic
PT   plants comprising resistance to biotic and abiotic stress.
XX
PS   Claim 2; Page 9; 77pp; English.
XX
CC   This sequence encodes a glutathione-S-transferase (GST) polypeptide,
CC   GST2, from the opium poppy. GST's catalyse the conjugation of
CC   electrophilic compounds to glutathione (GSH). They have been shown to be
CC   involved in detoxification of xenobiotics. The novel GST's of the
CC   invention GST1, GST2 and GST3, show extensive homology with tau type
CC   GST's from a variety of plant species, especially within the N-terminal
CC   region. This domain has been shown to be responsible for recognition of
CC   GSH and the ability of the proteins to form dimers. The C-terminal domain
CC   is responsible for substrate specificity and varies greatly between other
CC   GST's. The GST's of the invention exhibit strong glutathione conjugating
CC   activity towards a model substrate, 1-chloro-2,4-dinitrobenzene that
CC   could be inhibited in the presence of hydroxycinnamic acid amines of
CC   tyramine. The expression of the GST polypeptide in a transgenic plant is
CC   useful to produce a plant which has increased stress resistance. The
CC   plant has increased resistance to pathogens, herbicides or grazing pests.
CC   The GST polypeptide is useful to identify binding agents inhibitors and
CC   substrates
XX
SQ   Sequence 947 BP; 312 A; 135 C; 207 G; 293 T; 0 U; 0 Other;

  Query Match           29.2%;  Score 268.4;  DB 4;  Length 947;
  Best Local Similarity   63.5%;  Pred. No. 7.5e-58;
  Matches  410;  Conservative    0;  Mismatches  236;  Indels    0;  Gaps    0;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156

Db         48 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 107

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216

Db        108 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 167

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276

Db        168 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 227

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336

Db        228 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396

Db        288 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456

Db        348 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516

Db        408 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576

Db        468 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636

Db        528 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696

Db        588 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACGGTT 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742

Db        648 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 693


RESULT 10
ABZ14871
ID   ABZ14871 standard; DNA; 684 BP.
XX
AC   ABZ14871;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 2676.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-0227866P.
PR   26-JAN-2001; 2001US-0264647P.
PR   22-JUN-2001; 2001US-0300111P.
XX
```

PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 2676; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 684 BP; 184 A; 165 C; 169 G; 166 T; 0 U; 0 Other;

Query Match 26.1%; Score 240; DB 6; Length 684;
Best Local Similarity 59.6%; Pred. No. 1e-50;
Matches 405; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

```
Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
              |||||      |||| ||| |||  | ||  | |||   |||    |  | | |||||| |
Db          1 ATGGCGACCGAGGACGTGAAGCTGATCGGCTCATGGGCGAGTGTTTACGTCATGAGGGCG 60

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
               ||| || ||  |||||||   | |    || ||  |  || | |||| |    | |
Db         61 AGGATCGCTCTCCACCTCAAATCTATTAGCTACGAATTCCTTCAGGAGACGTATGGTTCA 120

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
              || || |||||||| || || || ||||| || |||||||| || || ||| |  | ||
Db        121 AAGAGCGAATTGCTCCTCAAATCAAACCCGGTTCACAAGAAGATGCCGGTTCTGATTCAC 180

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
              |  || |||    | || || ||    ||||| ||| | || || || |||| |||| |
Db        181 GCTGACAAACCGGTGTGCGAGTCCAACATCATCGTTCATTATATCGACGAGGCTTGGAAC 240

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
                 ||  |  ||||| || ||      || |||| || || || | ||| || ||||||
Db        241 TCTTCTGGACCTTCCATTCTCCCGTCCCATCCATACGACCGGGCCATTGCTCGGTTTTGG 300

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
              | | | ||||| || ||  ||||| |||  ||  ||| |||| | |||  |  || |||
Db        301 GCTGCCTACATAGACGATCAGTGGTTTATCTCTGTGAGAAGTATCCTAACAGCTCAAGGA 360

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
              || || |  |||||    |  | | |  ||||  ||||||  |     ||     ||||  ||
Db        361 GACGAAGAGAAGAAAGCAGCCATAGCTCAAGTTGAAGAAAGGACCAAGCTTCTGGAGAAA 420

Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
              |  |||||| | || ||  |||| ||  | |  |||   || ||    || || |  |
Db        421 GCATTCAACGATTGTAGCCAAGGAAAACCGTTCTTCAACGGTGACCATATCGGTTACCTC 480

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
              |||||||  || || ||||| ||  |||     ||||||   |||            |
Db        481 GACATTGCCTTGGGGAGCTTCTTAGGTTGGTGGAGAGTCGTCGAGTTGGATGCCAATCAT 540

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
              ||||| |||||||| || ||     ||   | |    |  |||||| ||     ||   || |
Db        541 AAATTTCTTGATGAGACCAAAACTCCCTCTCTAGTCAAATGGGCAGAGCGGTTCTGTGAT 600

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATT 744
              ||||| ||||||||      |  |||| |||| |  |||||| | ||| || |||| ||
```

Db 601 GATCCCGCTGTGAAACCTATAATGCCCGAGATTACAAAGCTCGCTGAATTCGCAAGGAAG 660

```
Qy        745 CTTCAGCTAAAATGGGCTGC 764
              ||    |  ||  ||   ||
Db        661 CTCTTTCCTAAGCGGCAAGC 680
```

RESULT 11
ABZ14853
ID ABZ14853 standard; DNA; 705 BP.
XX
AC ABZ14853;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2658.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 2658; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 705 BP; 204 A; 140 C; 170 G; 191 T; 0 U; 0 Other;

Query Match 25.3%; Score 232.2; DB 6; Length 705;
Best Local Similarity 59.7%; Pred. No. 1e-48;
Matches 409; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

```
Qy         92 AAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTG 151
              ||  |||  |||  || || || |  |||| ||||||||  ||| | ||     |||| |
Db         11 AAGAGGAAGTGAAACTATTAGGAGTATGGTACAGTCCATACGCCATAAGACCTAAGATCG 70

Qy        152 CCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATC---CCAAAA 208
              | |||   |||||     ||||||||||    ||||| || | | ||  |     |||| |
Db         71 CTCTTCGTCTCAAATCAGTGGATTATGATTACGTTGAGGAAAATCTGTTTGGATCCAAGA 130

Qy        209 GTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAG 268
              ||||  | ||||| || || ||||| || |||||||||  |||| ||| || ||||
Db        131 GTGAGCTTCTTCTCAAATCAAACCCGGTTCACAAGAAAGTCCCTGTTCTCCTCCACAACA 190
```

```
Qy        269 ATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACA 328
              | |||    ||     || |||    | ||| ||||| ||||| ||||||    |||  |
Db        191 ACAAACCGATTGTCGAGTCTCTCAACATCGTTGAATACATTGATGAGACGTGGAACTCAT 250

Qy        329 ATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTT 388
               ||| |  ||||| |||||    | || | |||||||| |||  ||| || ||||||  |
Db        251 CTGCACCGTCCATTCTTCCTTCACATCCTTATGATCGTGCTCTTGCTCGCTTTTGGTCTG 310

Qy        389 CTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATG 448
                | | | ||| |||||||| || || | |||| ||  |         ||| ||   || |
Db        311 ACTTCGTTGATAACAAGTGGTTTCCGGCGTTGAGAATGGCGGCGATCACTAAATCGGAAG 370

Qy        449 AGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGT 508
              | |||||    |  |  | | ||  ||| |||||||| | |    | |  | || || | ||
Db        371 ACGCAAAAGCGAAAGCCATGGAAGAAGTGGAAGAAGGGTTGTTGCAACTCGAGGATGCGT 430

Qy        509 TCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACA 568
              |          ||  |||||||  ||| |||||| || ||  | || || ||| | || |
Db        431 TTGTTTCTATAAGCAAAGGGAAACCCTTTTTCGGCGGTGAAGCAATCGGGTTTATGGATA 490

Qy        569 TTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAAT 628
              || | ||||||||||||| ||  | || | | ||    ||| ||  | || | |  |||
Db        491 TTTGCTTTGGAAGCTTTGTGGTTCTCTTGAAAGCTAGAGAAAAGTTTAAAGCAGAAAAGC 550

Qy        629 TGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATC 688
              |  | || ||| | ||     |||  | |       |||||| ||     ||       | |||
Db        551 TTTTAGACGAATCAAAAACTCCTTCTCTTTGTAAATGGGCCGACCGGTTCTTGTCCGATG 610

Qy        689 CTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTC 748
              | ||||||    |    || ||||  |||||  | | |||||||  | |    |||
Db        611 AAACGGTGAAGAATGTGGCACCGGAGATCGAAAAAGTAGCTGAGTTTCTACAAGAACTTG 670

Qy        749 AGCTAAAATGGGCTGCTGCAGCTGC 773
              || | | |       | ||||||
Db        671 AGGTTAGAGCTCAATCCGCAGCTTC 695
```

RESULT 12
ABN98975/c
ID ABN98975 standard; DNA; 817 BP.
XX
AC ABN98975;
XX
DT 01-AUG-2002 (first entry)
XX
DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 743.
XX
KW Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
KW disease; crop; thale cress; tolerance factor; insect; pathogen;
KW nutrition; ds.
XX
OS Arabidopsis thaliana.
XX
PN US2002023281-A1.
XX
PD 21-FEB-2002.
XX
PF 26-JAN-2001; 2001US-00770445.
XX
PR 27-JAN-2000; 2000US-0178472P.
XX
PA (GORL/) GORLACH J.
PA (ANYY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUYY/) YU Y.
PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHEW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
DR WPI; 2002-403163/43.
XX
PT New Arabidopsis thaliana nucleic acid for identifying homologous genes,
PT producing compositions that modulate the expression or function of its
PT encoded protein, and mapping functional regions of protein.
XX
PS Claim 1; SEQ ID NO 743; 49pp + Sequence Listing; English.
XX
CC The invention relates to an Arabidopsis thaliana nucleic acid (I)
CC comprising a sequence capable of hybridising under stringent conditions
CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
CC given in the specification or its fragment. A polypeptide (II) encoded by
CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
CC genetically modified cell (IV) comprising an exogenous nucleic acid, is
CC useful for screening a candidate agent for its biological effect. (I) is
CC useful in identifying homologous or related genes, in producing
CC compositions that modulate the expression or function of its encoded
CC protein, mapping functional regions of the protein and in studying
CC associated physiological pathways. (I) is also useful for the genetic
CC manipulation of cells, particularly plant cells. (I) is also useful in
CC screening assays of various plant strains to determine the strains that
CC are best capable of withstanding a particular disease or environmental
CC stress. (II) and (III) are useful for screening of biologically active
CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
CC pathways. The screened agents are useful in improved methods of treating
CC crops to prevent or treat disease. (II) are also useful in screening
CC programs to identify agents that mimic or enhance the action of tolerance
CC factors. Such agents are useful in improved methods of treating crops to
CC enhance their tolerance to environmental stress. (I) is also useful for
CC enhancing or inhibiting production of a biosynthetic product in a plant.
CC (III) is useful for identifying other mediators that may induce
CC expression of proteins of interest, for establishing the extent to which
CC any specific insect and/or pathogen is responsible for damage to a
CC particular plant, for identifying other mediators that enhance or induce
CC tolerance to environmental stress, for identifying factors involved in
CC biosynthetic pathways of nutritional, commercial, or medicinal value and
CC for identifying productions of nutritional, commercial or medicinal
CC value. (IV) is useful in the study of genetic function and regulation,
CC for alteration of the cellular metabolism and for screening compounds
CC that may affect the biological function of the gene or gene products.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=999909770445
XX
SQ Sequence 817 BP; 233 A; 190 C; 130 G; 250 T; 0 U; 14 Other;

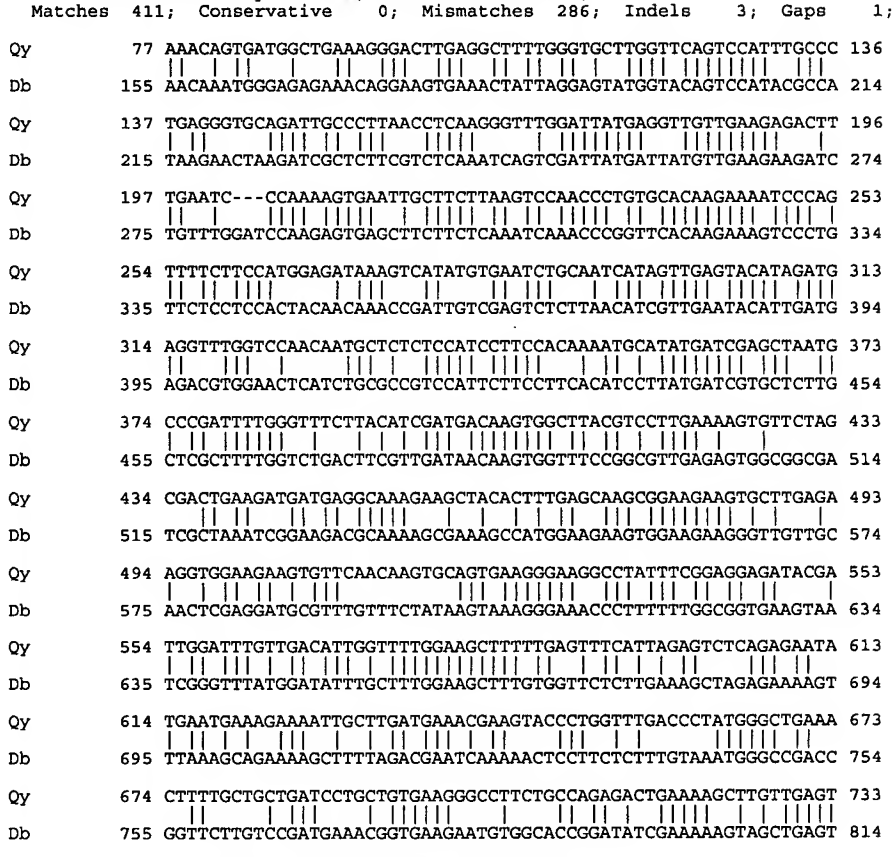
Query Match 24.9%; Score 228.2; DB 6; Length 817;
Best Local Similarity 59.3%; Pred. No. 1.1e-47;
Matches 396; Conservative 0; Mismatches 269; Indels 3; Gaps 1;

```
Qy         70 TTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCA 129
                ||| |  |||||  |||||  || || ||  |||  || |||||  | |||| ||||||
Db        817 TTAGACAAAACAAAGATGGTAGAGAGAGAGGTGAAACTGTTGGGAACATGGTACAGTCCG 758

Qy        130 TTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAA 189
                | |   | || ||  |||| || |||   |||||     | ||||||||    ||||||
Db        757 GTCGTGATAAGAGTAAAGATCGCTCTTCGTCTCAAATCGGTTGATTATGATTACGTTGAA 698

Qy        190 GAGACTTTG---AATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAA 246
```

```
Db        697 GAGAACTTGTTCGGATCTAAGAGTGAGCTTCTTCTCAAATCGAACCCGGTTTACAAGAAA 638

Qy        247 ATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTAC 306

Db        637 GTGCCGGTTCTCATCCACAGTAATAAATCGGTTTGTGAGTCTCTAAACATCGTTGAATAC 578

Qy        307 ATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGA 366

Db        577 ATAGACGAGACGTGGAACTCATCTGGATCGTTCATTCTTCCTTCTCATCCTCATGATCGT 518

Qy        367 GCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGT 426

Db        517 GCCCTTGCTCGCTTCTGGTCTGCCTTCGTTGATGACAAGTGGTTTCCGGCTTTGATGGCA 458

Qy        427 GTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTG 486

Db        457 GCTGTGGTCGCTAAATCGGAAGACGCAAAAGCAAAAGGCATGGAAGAAGTGGAAGAAGGG 398

Qy        487 CTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGA 546

Db        397 TTGTTGCAACTCGAAGCTGCGTTTATTGCTCTAAGCAAAGGAAAATCTTTTTTCGGCGGT 338

Qy        547 GATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCA 606

Db        337 GAAACAATCGGGTTCATAGACATTTGCCTTGGAAGCTTTTTGGTTTTCTTGAAAGCTAGA 278

Qy        607 GAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGG 666

Db        277 GAGAAGCTTAAAAAAGNNNNNATTNTAGACGAATNNAAAACTCCTTCTCTTTATAGATGG 218

Qy        667 GCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTT 726

Db        217 GCCAACCGNNNTTTGTCAGATGAAACAGTGAAGAATGTGGTGCCGGAGATAGATAAAGTT 158

Qy        727 GTTGAGTT 734

Db        157 GCCAAGCT 150


RESULT 13
AAC41255
ID   AAC41255 standard; DNA; 955 BP.
XX
AC   AAC41255;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 31218.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
```

PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

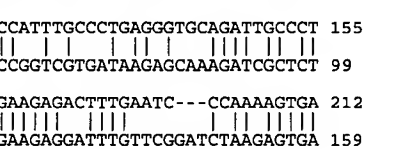
Query Match 24.7%; Score 226.4; DB 3; Length 955;
Best Local Similarity 58.7%; Pred. No. 3.3e-47;
Matches 411; Conservative 0; Mismatches 286; Indels 3; Gaps 1;



```
                ||  | |    ||| || | | |       | ||||||
Db        815 TTCTACAAGAACTTGAGGTTAGAGCTCAATCCGCAGCTTC 854

RESULT 14
AAC51154
ID   AAC51154 standard; DNA; 863 BP.
XX
AC   AAC51154;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 67471.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
```

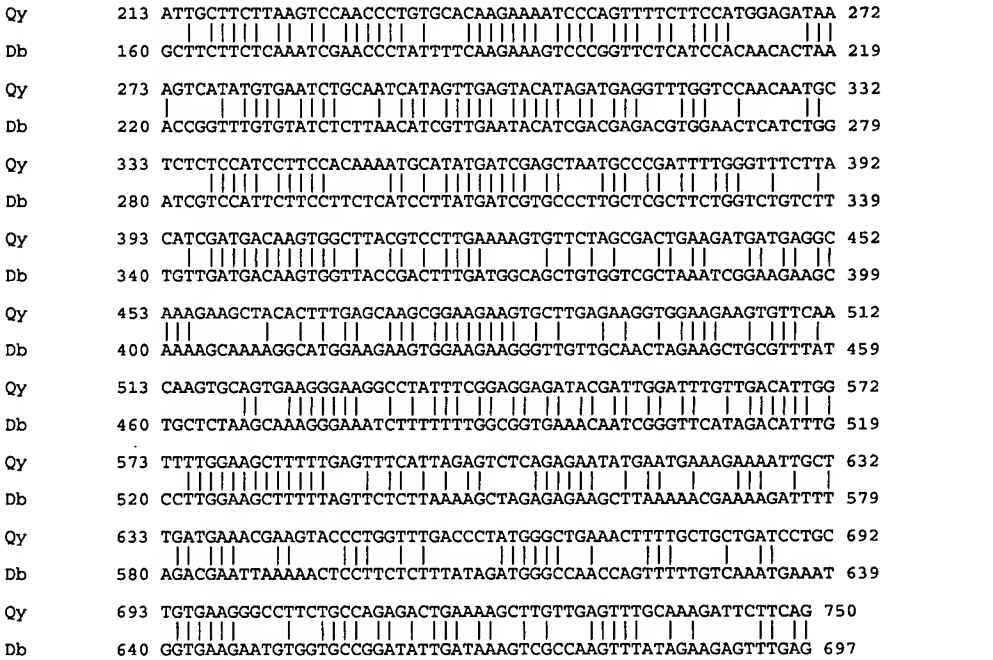
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 22.0%; Score 202.2; DB 3; Length 863;
Best Local Similarity 57.8%; Pred. No. 4.3e-41;
Matches 380; Conservative 0; Mismatches 275; Indels 3; Gaps 1;



Qy 96 GGACTTGAGGCTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCT 155
Db 40 GGAAGTGAAACTGTTGGGAACATGGTACAGTCCGGTCGTGATAAGAGCAAAGATCGCTCT 99

Qy 156 TAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATC---CCAAAAGTGA 212
Db 100 TCGTCTCAAATCGGTTGATTATGATTACGTTGAAGAGGATTTGTTCGGATCTAAGAGTGA 159



Qy 213 ATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAA 272
Db 160 GCTTCTTCTCAAATCGAACCCTATTTTCAAGAAAGTCCCGGTTCTCATCCACAACACTAA 219

Qy 273 AGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGC 332
Db 220 ACCGGTTTGTGTATCTCTTAACATCGTTGAATACATCGACGAGACGTGGAACTCATCTGG 279

Qy 333 TCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTA 392
Db 280 ATCGTCCATTCTTCCTTCTCATCCTTATGATCGTGCCCTTGCTCGCTTCTGGTCTGTCTT 339

Qy 393 CATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGC 452
Db 340 TGTTGATGACAAGTGGTTACCGACTTTGATGGCAGCTGTGGTCGCTAAATCGGAAGAAGC 399

Qy 453 AAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAA 512
Db 400 AAAAGCAAAAGGCATGGAAGAAGTGGAAGAAGGGTTGTTGCAACTAGAAGCTGCGTTTAT 459

Qy 513 CAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGG 572
Db 460 TGCTCTAAGCAAAGGGAAATCTTTTTTTGGCGGTGAAACAATCGGGTTCATAGACATTTG 519

Qy 573 TTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCT 632
Db 520 CCTTGGAAGCTTTTTAGTTCTCTTAAAAGCTAGAGAGAAGCTTAAAAACGAAAAGATTTT 579

Qy 633 TGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGC 692
Db 580 AGACGAATTAAAAACTCCTTCTCTTTATAGATGGGCCAACCAGTTTTTGTCAAATGAAAT 639

Qy 693 TGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAG 750
Db 640 GGTGAAGAATGTGGTGCCGGATATTGATAAAGTCGCCAAGTTTATAGAAGAGTTTGAG 697

RESULT 15
AAC35433
ID AAC35433 standard; DNA; 866 BP.
XX
AC AAC35433;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 10180.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.

```
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             21.5%;  Score 197.4;  DB 3;  Length 866;
  Best Local Similarity   57.3%;  Pred. No. 7.1e-40;
  Matches  377;  Conservative    0;  Mismatches  278;  Indels    3;  Gaps    1;

Qy          96 GGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCT 155
               | |  |||  || |||||  | |||| ||||||  | |   | || |   |||| |  ||
Db          40 GCAAGTGAAACTGTTGGGAACATGGTACAGTCCGGTCGTGATAAGAGCAAAGATCGATCT 99

Qy         156 TAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATC---CCAAAAGTGA 212
               |  |||||      | ||||||||    ||||||||||  |||        | || |||||
Db         100 TCGTCTCAAATCGGTTGATTATGATTACGTTGAAGAGAACTTGTTCGGATCTAAGAGTGA 159

Qy         213 ATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAA 272
               | ||||| || || ||||| ||  |||||||| |||| ||| || ||||       |||
Db         160 GCTTCTTCTCAAATCGAACCCGGTTTACAAGAAAGTCCCGGTTCTCATCCACAACACTAA 219

Qy         273 AGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGC 332
               |    | ||||  |||   | ||| ||||| ||||| || |||   |||  |     ||
Db         220 ACCGGTTTGTGTGTCTCTTAACATCGTTGAATACATCGACGAGACGTGGAACTCATCTGG 279

Qy         333 TCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTA 392
               |||||  |||||      || | |||||||| ||   ||| || || |||  |   |
Db         280 ATCGTCCATTCTTCCTTCTCATCCTTATGATCGTGCCCTTGCTCGCTTCTGGTCTGTCTT 339

Qy         393 CATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGC 452
               | ||||||||||||  |  || | ||||     | | | |   || ||   || || ||
Db         340 TGTTGATGACAAGTGGTTACCGACTTTGATGGCAGCTGTGGTCGCTAAATCGGAAGAAGC 399

Qy         453 AAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAA 512
               |||       |  | | ||  ||| |||||||| | |   | |  | || |  | ||| |
Db         400 AAAAGCAAAAGGCATGGAAGAAGTGGAAGAAGGGTTGTTGCAACTAGATGCTGCGTTTAT 459

Qy         513 CAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGG 572
                  ||  |||||||  | | ||| || || || || || || ||  | ||||| |
Db         460 TGCTCTAAGCAAAGGGAAATCTTTTTTTGGCGGTGAAACAATCGGGTTCATAGACATTTG 519

Qy         573 TTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCT 632
               ||||||||||||   | || | | ||    ||||||  | ||  |   |||  | |
Db         520 CCTTGGAAGCTTTTTAGTTCTCTTAAAAGCTAGAGAGAAGCTTAAAAACGAAAAGATTTT 579

Qy         633 TGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGC 692
               || |||   ||    |||  | |       ||||||  |    |||    |  ||
Db         580 AGACGAATTAAAAACTCCTTCTCTTTATAGATGGGCCAACCAGTTTTTGTCAAATGAAAT 639

Qy         693 TGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAG 750
               ||||||    | |||| || | ||| ||  | |   |||||  | |     || ||
Db         640 GGTGAAGAATGTGGTGCCGGATATTGATAAAGTCGCCAAGTTTATAGAAGAGTTTGAG 697


Search completed: October 13, 2004, 03:53:58
Job time : 451 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 13, 2004, 03:46:31 ; Search time 89 Seconds
(without alignments)
5724.100 Million cell updates/sec

Title: US-10-088-945A-14
Perfect score: 918
Sequence: 1 cataaaactccacatttcct.....aatatagtaaaaaaaaaaaa 918

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	620.2	67.6	895	3	US-08-924-747-7	Sequence 7, Appli
2	620.2	67.6	895	3	US-09-247-373B-7	Sequence 7, Appli
3	620.2	67.6	895	3	US-09-296-715-7	Sequence 7, Appli
4	167.4	18.2	970	3	US-09-248-335-47	Sequence 47, Appl
5	166.2	18.1	895	2	US-08-924-759-23	Sequence 23, Appl
6	166.2	18.1	895	3	US-09-248-335-23	Sequence 23, Appl
7	164	17.9	1068	3	US-09-248-335-73	Sequence 73, Appl
8	157.6	17.2	904	3	US-09-248-335-69	Sequence 69, Appl
9	156	17.0	960	3	US-09-248-335-57	Sequence 57, Appl
10	152.4	16.6	934	3	US-09-248-335-55	Sequence 55, Appl
11	142.6	15.5	900	3	US-09-248-335-61	Sequence 61, Appl
12	140	15.3	841	3	US-09-247-373B-53	Sequence 53, Appl
13	139.8	15.2	860	3	US-09-248-335-41	Sequence 41, Appl
14	131.4	14.3	1013	3	US-09-248-335-71	Sequence 71, Appl
15	130.8	14.2	872	3	US-09-248-335-63	Sequence 63, Appl
16	130.8	14.2	931	3	US-08-924-747-9	Sequence 9, Appli
17	130.8	14.2	931	3	US-09-247-373B-9	Sequence 9, Appli
18	130.8	14.2	931	3	US-09-296-715-9	Sequence 9, Appli
19	129.8	14.1	967	3	US-09-248-335-59	Sequence 59, Appl
20	129.6	14.1	1074	3	US-09-248-335-67	Sequence 67, Appl
21	128.8	14.0	859	3	US-09-247-373B-47	Sequence 47, Appl
22	128.2	14.0	885	3	US-08-924-747-23	Sequence 23, Appl
23	128.2	14.0	885	3	US-09-247-373B-23	Sequence 23, Appl
24	128.2	14.0	885	3	US-09-296-715-23	Sequence 23, Appl
25	128.2	14.0	1100	3	US-09-248-335-53	Sequence 53, Appl
26	123.2	13.4	977	3	US-08-924-747-13	Sequence 13, Appl
27	123.2	13.4	977	3	US-09-247-373B-13	Sequence 13, Appl
28	123.2	13.4	977	3	US-09-296-715-13	Sequence 13, Appl
29	123	13.4	895	3	US-08-924-747-21	Sequence 21, Appl
30	123	13.4	895	3	US-09-247-373B-21	Sequence 21, Appl
31	123	13.4	895	3	US-09-296-715-21	Sequence 21, Appl
32	119	13.0	836	3	US-09-247-373B-49	Sequence 49, Appl
33	118.8	12.9	993	3	US-08-924-747-17	Sequence 17, Appl
34	118.8	12.9	993	3	US-09-247-373B-17	Sequence 17, Appl
35	118.8	12.9	993	3	US-09-296-715-17	Sequence 17, Appl
36	117.4	12.8	1083	3	US-09-247-373B-35	Sequence 35, Appl
37	116.6	12.7	1117	3	US-09-247-373B-33	Sequence 33, Appl
38	114.8	12.5	853	3	US-09-247-373B-51	Sequence 51, Appl
39	114.6	12.5	1006	3	US-08-924-747-15	Sequence 15, Appl
40	114.6	12.5	1006	3	US-09-247-373B-15	Sequence 15, Appl
41	114.6	12.5	1006	3	US-09-296-715-15	Sequence 15, Appl
42	110.6	12.0	1179	3	US-09-247-373B-43	Sequence 43, Appl
43	109.4	11.9	991	3	US-08-924-747-25	Sequence 25, Appl
44	109.4	11.9	991	3	US-09-247-373B-25	Sequence 25, Appl
45	109.4	11.9	991	3	US-09-296-715-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-08-924-747-7
; Sequence 7, Application US/08924747
; Patent No. 6063570
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; TITLE OF INVENTION: ENZYMES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,747
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 895 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SE3.03B09
US-08-924-747-7

Query Match 67.6%; Score 620.2; DB 3; Length 895;
Best Local Similarity 88.0%; Pred. No. 2.8e-164;

```
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| |||||||||  |  ||||||| ||||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || ||||||||||||||||||| ||||||||||||||||   ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||||| |||||||| |||||||| |||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | |||||||||||||||||||||||||||||||| ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||| ||||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||| |||||||||||||||||| ||||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || |||||||||||||||||||||||||||||  || | |
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  |||||||||| |||||| ||| ||||| |||||||| |||||||||||||||  |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |     | |   | || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 2
US-09-247-373B-7
; Sequence 7, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
;  APPLICANT:  MCGONIGLE, BRIAN
;  APPLICANT:  O'KEEFE, DANIEL
;  TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE:  CL-1108-A
;  CURRENT APPLICATION NUMBER: US/09/247,373B
;  CURRENT FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  08/924,747
;  PRIOR FILING DATE:  1997-09-05
;  NUMBER OF SEQ ID NOS:  56
;  SOFTWARE:  Microsoft Office 97
; SEQ ID NO 7
;   LENGTH: 895
;   TYPE: DNA
;   ORGANISM: SOYBEAN
US-09-247-373B-7

  Query Match             67.6%;  Score 620.2;  DB 3;  Length 895;
  Best Local Similarity   88.0%;  Pred. No. 2.8e-164;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| |||||||||  |  ||||||| ||||||||||||||||||||||||||||||
Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || ||||||||||||||||||| ||||||||||||||||   ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||||| |||||||| |||||||| |||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | |||||||||||||||||||||||||||||||| ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
              |||| |||| |||||||||||||| |||||||||| ||||||||||||||||||||||||
Db        588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
              |||||| |||| |||||||||| |||||||||||||||||| ||||||||||||||||||||
Db        648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy        757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
              ||| ||||||||      || |||||||||||||||||||||||||||||  || | |
Db        708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy        811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
              |||  |||||||||| |||||| ||| ||||| |||||||| |||||||||||||||  |
Db        767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy        871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              |||||||||| |   ||||| | |     | |   | || |||
Db        826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 3
US-09-296-715-7
; Sequence 7, Application US/09296715
; Patent No. 6171839
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/296,715
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  895 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  SOYBEAN
;    IMMEDIATE SOURCE:
;      CLONE:  SE3.03B09
US-09-296-715-7

  Query Match             67.6%;  Score 620.2;  DB 3;  Length 895;
  Best Local Similarity   88.0%;  Pred. No. 2.8e-164;
  Matches  724;  Conservative    0;  Mismatches   88;  Indels   11;  Gaps    4;

Qy          97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
               ||||||| |||||||||  |  ||||||| ||||||||||||||||||||||||||||||
Db          51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy         157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
               |||||||||||| | || |||||||||||||||||||| ||||||||||||||||  ||
Db         111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 170

Qy         217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy         277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
               || |||||||||||||||||||||||||||||||| |||||||  ||||   |  |||| || |
Db         231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy         337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
               |||||||||||||||||| |||||||| |||||||| ||||||||| || | |||||
Db         288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy         397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
               ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db         348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy         457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
               |||| |||||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db         408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy         517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
               | |||||||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db         468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy         577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
               || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||||||||
Db         528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 587

Qy         637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
               |||| |||| ||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db         588 GAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTG 647

Qy         697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAA 756
               |||||| |||| |||||||||||| |||||||||||||||||| ||||||||||||||||||
Db         648 AAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAA 707

Qy         757 TGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAA 810
               ||| ||||||||      || ||||||||||||||||||||||||||||||  || | |
Db         708 TGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTA 766

Qy         811 TTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATAT 870
               |||  |||||||||||| |||||| ||| ||||| |||||||| ||||||||||||||| |
Db         767 TTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTT 825

Qy         871 ATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
               |||||||||| |  ||||| | |     | |   | || |||
Db         826 ATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 868


RESULT 4
US-09-248-335-47
; Sequence 47, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 47
;   LENGTH: 970
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-47

  Query Match             18.2%;  Score 167.4;  DB 3;  Length 970;
  Best Local Similarity   52.7%;  Pred. No. 2.7e-37;
  Matches  363;  Conservative    0;  Mismatches  326;  Indels    0;  Gaps    0;

Qy          95 GGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCC 154
               || |  ||| |||  |||| || |||   || || || |     | || |  || || |
Db          90 GGCATCTGAAGCTGCTGGGCGCGTGGCCGAGCCCCTTCGTGAACCGTGTACGCATGGCGC 149

Qy         155 TTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAAT 214
               |  ||||||||||  |||| || |||   || || |||    | |    ||| || ||
Db         150 TGCACCTCAAGGGGCTGGAGTACGAGAACGTGGAGGAGGACCTCACCAACAAGAGCGACC 209

Qy         215 TGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAG 274
               |||| ||    |||||||| || ||||||    | || ||  || |||| || || ||
Db         210 TGCTCCTCGCCTCCAACCCCGTCCACAAGCTCGTTCCCGTCCTCCTCCACGGCGACAAGC 269

Qy         275 TCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTC 334
               |||  |  || ||    |||| || |||||| | || || |  |   ||  |  || |
Db         270 CCATCTCCGAGTCACTCGTCATCGTGGAGTACCTCGACGACGCCTTCCCCGGCGCTGGCC 329

Qy         335 TCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACA 394
```

```
Db        330 AGGCCGTCCTCCCCGCCGACCCCTATGAACGCGCCGTCGCTCGCTTCTGGGCCAAATACG 389

Qy        395 TCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAA 454

Db        390 TCGACGGCAAGTTGCACGGCATGATGGTGAAGGCGCTCATGGGGGCAACGGAGGAGGAGA 449

Qy        455 AGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACA 514

Db        450 GGGCGACGGCGACGGTGGACGCGCTGGCCGCTATGGACACGCTGGAGGGCGCGTTCGCCG 509

Qy        515 AGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTT 574

Db        510 AGTGCTCCGGCGGGAAAAAGTTCTTCGCCGGCGACGCGCCCGGGTACCTGGACGTCGCGC 569

Qy        575 TTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTG 634

Db        570 TGGGAGGCTTCATCGGCTGGCTGCGCGCCTGGGACAAGGTGGGGGGCGTCAAGCTGCTGG 629

Qy        635 ATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTG 694

Db        630 ACGCCGGCCGGGTCCCGCGCCTGGCCACGTGGGCGGAGCGCTTCGCCGCGCTCGACGTAG 689

Qy        695 TGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAA 754

Db        690 CCAAGGAGGTCATCCCGGACCCCGACCACATCGCCGAGTTTGCCAAGGTGCTGCAGGCAC 749

Qy        755 AATGGGCTGCTGCAGCTGCTGCAAAGTAA 783

Db        750 GCTCCGCGGCAGCTGCCACCAGCAACTGA 778


RESULT 5
US-08-924-759-23
; Sequence 23, Application US/08924759
; Patent No. 5962229
;  GENERAL INFORMATION:
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  PLANT GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/924,759
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1128
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  895 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  MAIZE
;    IMMEDIATE SOURCE:
;      CLONE:  CS1.PK0059.E2
US-08-924-759-23

  Query Match             18.1%;  Score 166.2;  DB 2;  Length 895;
  Best Local Similarity   53.2%;  Pred. No. 5.7e-37;
  Matches  374;  Conservative    0;  Mismatches  328;  Indels    1;  Gaps    1;

Qy         72 ATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATT 131

Db         39 AGAGTCTATAATGGCGGACGGAGGCGAGCTGCAGCTGCTGGGCTCATGGTACAGCCCCTA 98

Qy        132 TGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGA 191

Db         99 CGTGATCCGCGCCAAGGTGGCGCTGGGGCTGAAGGGGCTCAGCTACGAGTTCGTCGAGGA 158

Qy        192 GACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCC 251

Db        159 GGACCTCTCCCGCAAGAGCGACCTGCTGCTGAAGCTCAACCCGGTGCACAGGAAGGTGCC 218

Qy        252 AGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGA 311

Db        219 CGTGCTGGTCCACGGCGGCCGCCCCGTGTGCGAGTCGCTCGTCATCCTGCAGTACGTCGA 278

Qy        312 TGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAA 371

Db        279 CGAGACCTGGGCAGGCACCGGGACCCCTCTCCTCCCCGCCGACGCCTACGACCGCGCCAT 338

Qy        372 TGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCT 431

Db        339 GGCTCGCTTCTGGGCAGCCTACGTCGACGACAAGTTCTACAAGGAGTGGAACCGGCTGTT 398

Qy        432 AGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGA 491

Db        399 CTGGTCGACGACGGCGGAGAAGGCGGCGGAGGCGCTCGGCGTCGTCGTCCCCGTGGTGGA 458

Qy        492 GAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATAC 551

Db        459 GACGCTGGAGCAGGCGTTCAGGGAGTGCTCCAAAGGGAA-ACCTTCTTCGGCGGCGACGC 517

Qy        552 GATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAA 611

Db        518 CGTCGGGCTCGTGGACATCGCGCTCGGGAGCTTCGTGGTGTGGATCAGGGTGGTGGACGA 577

Qy        612 TATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGA 671

Db        578 GGCGGCCGGCGTAAAGCTTCTGGACGAGGCCAAGTTCCCGGCCTTGACGGCGTGGGCGGA 637

Qy        672 AACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGA 731

Db        638 GCGCTTCTTGGCGGTGGACGCCGTGAAGGAGGTGATGCCGGACGCCGGAAGGCTGTTGGA 697

Qy        732 GTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCT 774

Db        698 GCACTACAAGGGGTTTCTGGCTAAACGGTCTCCACCTGCTGGT 740


RESULT 6
US-09-248-335-23
; Sequence 23, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
```

```
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 23
;   LENGTH: 895
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-23

  Query Match             18.1%;  Score 166.2;  DB 3;  Length 895;
  Best Local Similarity   53.2%;  Pred. No. 5.7e-37;
  Matches  374;  Conservative    0;  Mismatches  328;  Indels    1;  Gaps    1;

Qy         72 ATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATT 131
Db         39 AGAGTCTATAATGGCGGACGGAGGCGAGCTGCAGCTGCTGGGCTCATGGTACAGCCCCTA 98

Qy        132 TGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGA 191
Db         99 CGTGATCCGCGCCAAGGTGGCGCTGGGGCTGAAGGGGCTCAGCTACGAGTTCGTCGAGGA 158

Qy        192 GACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCC 251
Db        159 GGACCTCTCCCGCAAGAGCGACCTGCTGCTGAAGCTCAACCCGGTGCACAGGAAGGTGCC 218

Qy        252 AGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGA 311
Db        219 CGTGCTGGTCCACGGCGGCCGCCCCGTGTGCGAGTCGCTCGTCATCCTGCAGTACGTCGA 278

Qy        312 TGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAA 371
Db        279 CGAGACCTGGGCAGGCACCGGGACCCCTCTCCTCCCCGCCGACGCCTACGACCGCGCCAT 338

Qy        372 TGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCT 431
Db        339 GGCTCGCTTCTGGGCAGCCTACGTCGACGACAAGTTCTACAAGGAGTGGAACCGGCTGTT 398

Qy        432 AGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGA 491
Db        399 CTGGTCGACGACGGCGGAGAAGGCGGCGGAGGCGCTCGGCGTCGTCGTCCCCGTGGTGGA 458

Qy        492 GAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATAC 551
Db        459 GACGCTGGAGCAGGCGTTCAGGGAGTGCTCCAAAGGGAA-ACCTTCTTCGGCGGCGACGC 517

Qy        552 GATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAA 611
Db        518 CGTCGGGCTCGTGGACATCGCGCTCGGGAGCTTCGTGGTGTGGATCAGGGTGGTGGACGA 577

Qy        612 TATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGA 671
Db        578 GGCGGCCGGCGTAAAGCTTCTGGACGAGGCCAAGTTCCCGGCCTTGACGGCGTGGGCGGA 637

Qy        672 AACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGA 731
Db        638 GCGCTTCTTGGCGGTGGACGCCGTGAAGGAGGTGATGCCGGACGCCGGAAGGCTGTTGGA 697

Qy        732 GTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCT 774
Db        698 GCACTACAAGGGGTTTCTGGCTAAACGGTCTCCACCTGCTGGT 740
```

```
RESULT 7
US-09-248-335-73
; Sequence 73, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 73
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-73

  Query Match             17.9%;  Score 164;  DB 3;  Length 1068;
  Best Local Similarity   53.1%;  Pred. No. 2.6e-36;
  Matches  372;  Conservative    0;  Mismatches  325;  Indels    3;  Gaps    1;

Qy         88 GCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAG 147
Db         35 GGTGACGATGAACTCAAGCTGCTGGGGGCGTGGGCGAGCCCATTCGTCCTGCGGGTGAAG 94

Qy        148 ATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCC---C 204
Db         95 CTCGCGCTCAGCTTCAAGGGCCTGAGCTACGAGGACGTGGAGGAGGACCTCTCCGGCGGC 154

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
Db        155 AAGAGCGAGCTGCTCCTCGAGTCCAACCCGGTGCACAAGAAGGTGCCCGTGCTCCTCCAC 214

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
Db        215 AACGGCAAGCCTGTGTGCGAGTCGCAGATCATCGTGCAGTACATCGATGAGGCCTTCGCC 274

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
Db        275 GGCACTGGCCCGTCCCTTCTCCCTGCCGACCCGCACCAGCGCGCCGTGGCTCGCTTCTGG 334

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
Db        335 GGTGCCTACATTGACGACAAGCTCCTAGCCTTCTGGCTGCAATCAGCAAGGGCCAAGACG 394

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
Db        395 CAGGAGGAAAAGGCCGAGGCGCTGAAGCAGGCGCTCGCCGCGGCCGAGAACCTGGAGGCC 454

Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
Db        455 GCCTTCACGGAGATCTCCGAGGGCAAGCCCTTCTTCGGCGGCGACAGCGTCGGGTACCTG 514

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
Db        515 GACGTGACGCTGGGAGCGCTGGTCGCGTGGGTGCACGCCGCCGAGAAGCTGTACGGGATG 574

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
Db        575 AGGCTCTTCGACGCCACGAGGACCCCGCGGCTGAGCGCGTTCGTGGAGAGGTTCGGCGCG 634

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATT 744
Db        635 CTCGGAGCGGCCAAGGCGGTGCTGCCCGACGTCGATGGCCTCGTCGAATACGCCAAACAG 694

Qy        745 CTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAGTAAA 784
Db        695 AGGCAGGCCGACGCGGCAGCTGCAGCCTCGGACAGCTAAA 734
```

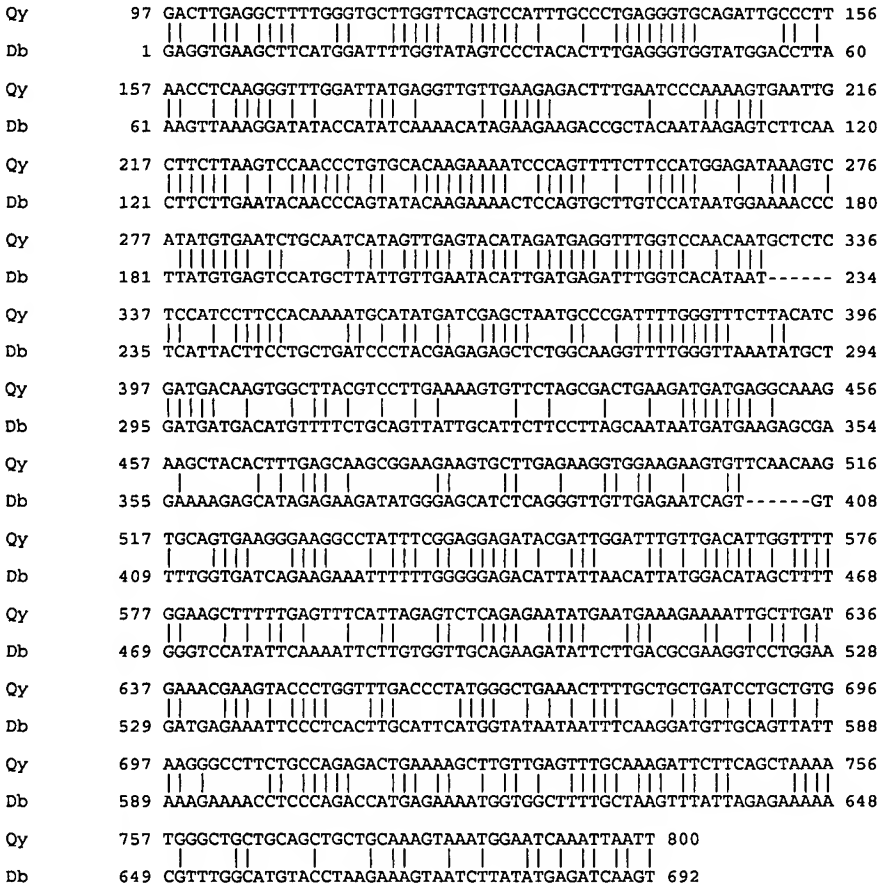
```
RESULT 8
US-09-248-335-69
; Sequence 69, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
```

```
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 69
;   LENGTH: 904
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-69

  Query Match            17.2%;  Score 157.6;  DB 3;  Length 904;
  Best Local Similarity  54.5%;  Pred. No. 1.5e-34;
  Matches  338;  Conservative    0;  Mismatches  279;  Indels    3;  Gaps    1;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
Db         85 GAGCTGAAGCTGCTGGGCGTGTGGGACAGCCCGTACGTCAACAGGGTCCAGATCGTGCTC 144

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
Db        145 AACCTCAAGGGCCTCAGCTACGAGTACGTGGAGGAGGACCTCCTCAGCAAGAGCGAGCTC 204

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
Db        205 CTCCTCAATTCCAACCCGGTGCACAAGAAAGTGCCCGTGCTCATCCACGCCGGCAAGCCG 264

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
Db        265 GTCGCCGAGTCGCAGGCCATCGTTCAGTACCTCGACGAGGCTTTCCCCAGCGGCACGTTC 324

Qy        337 ---TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTAC 393
Db        325 CCGTCGGTCCTCCCAGCCGAACCCTACGCACGCGCCACCGCCCGCTTCTGGGCCGCCTTC 384

Qy        394 ATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCA 453
Db        385 GTCGACGACAAGGTCGGGTCTCCATGGCACACGGTCCTGTTCGCGCGGGAGCACGGGAAG 444

Qy        454 AAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAAC 513
Db        445 AAGGCGGACGCGGCGTCGCGGATCGTCGCGGCGCTGGAGACGCTGGAGGGTGCGTTCGAG 504

Qy        514 AAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGT 573
Db        505 GACTGCTCCGGCGGGAGGGACTACTTCGGCGGCGACGCCATCGGCTTCGTGGACGTGGTC 564

Qy        574 TTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTT 633
Db        565 CTCGGCAGCTACCTGGGCTGGTTCAAGGTGTTCGAGAAGATGGTCGGCGTCAGGGTCCTG 624

Qy        634 GATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCT 693
Db        625 GACGTGGCGAGGACGCCGCTCCTCGCCGCGTGGGGGGAGCGTTTCGCGGCGGCGGAAGCG 684

Qy        694 GTGAAGGGCCTTCTGCCAGA 713
Db        685 GCCAAGGACGTCCTGCCGGA 704

RESULT 9
US-09-248-335-57
; Sequence 57, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 57
;   LENGTH: 960
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-57

  Query Match            17.0%;  Score 156;  DB 3;  Length 960;
  Best Local Similarity  52.6%;  Pred. No. 4.3e-34;
  Matches  362;  Conservative    0;  Mismatches  325;  Indels    1;  Gaps    1;

Qy         92 AAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTG 151
Db          1 AACCGCAGCTGAAGCTGCTGGCCATGTGGGCGAGCCCGTTTGCCCTACGGGCGAAGCTAG 60

Qy        152 CCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTG 211
Db         61 CGCTCAACTTCAAGGGCCTGGCCTACGAGTACGTAGAGGAGGACCTCCGCAGCAAGAGCG 120

Qy        212 AATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATA 271
Db        121 ACCTCCTGCTGAGCTCGAACCCGGTGCACAAGAAGGTGCCCGTCCTCATCCACAACGGCG 180

Qy        272 AAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATG 331
Db        181 TGCCCGTCTGTGAGTCGCGGGTCATCGTGGAGTACCTCGACGAAGTCTACAGCGCCACGG 240

Qy        332 CTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTT 391
Db        241 GCCCCCGCTTCCTCCCTGCCGACCCATACGAGCGTGCCATGGCGCGCTTCTGGGCCTCAT 300

Qy        392 ACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGG 451
Db        301 TCATCGACGAAAAGTTCTTGGCGTCGTGGCTAAAGGCAGGAAGGGGCAAGACGGACGAGG 360

Qy        452 CAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCA 511
Db        361 AGAAGGCCGAAGGGTTGAAGCTGACACTCGCGGCCGTAGAAACCTTGGAAGGGGCGTTCA 420

Qy        512 ACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTG 571
Db        421 TGGAGTGCTCCAAGGGGAAGCCCTTCTTTGGAGGCGATAGTGTCGGCTACCTGGACATCG 480

Qy        572 GTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGC 631
Db        481 CGCTCGGGGCCCTGGTAGCGTGGATGCGCGCCACCGAGGCGCGTCATGGTCTCAGGCTCT 540

Qy        632 TTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTG 691
Db        541 TCGACGCCTCCAGAGTCCGCTGCTGGAGAA-GTGGGTGGAGCGCTTCAGCGAGCTGGACG 599

Qy        692 CTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGC 751
Db        600 AGGTCGTGGCGGTCATGCCGGACATCGACCGGCTAGTAGAGCTCGGCAAGGTGAGGGAGG 659

Qy        752 TAAAATGGGCTGCTGCAGCTGCTGCAAA 779
Db        660 CTGCTGCGGCTGCAGCAGCTGCCGTAAA 687

RESULT 10
US-09-248-335-55
; Sequence 55, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
```

```
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 55
;   LENGTH: 934
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-55

  Query Match            16.6%;  Score 152.4;  DB 3;  Length 934;
  Best Local Similarity  53.2%;  Pred. No. 4.3e-33;
  Matches  346;  Conservative    0;  Mismatches  301;  Indels    3;  Gaps    1;

Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
              ||| | ||    | | || |  |  | ||    |||   || || || |   |  | ||
Db         54 ATGTCAGAGGCCGCCGTGCGTGTGATCGGCCTATGGCCGAGCCCGTTCGTGATCCGCGTC 113

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
              | ||| ||||| || || |||||  | || |  ||| | || || |||     ||
Db        114 CTGATCGCCCTGAAGCTGAAGGGCGTCGAGTTCGAGTTCGTGGAGGAGGTGGTGGGCAGG 173

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
              || || ||  |||| || | ||| ||||| |||||||||||| ||||| ||   |  ||||
Db        174 AAGAGCGAGCTGCTGCTGAGGTCGAACCCGGTGCACAAGAAGATCCCCGTCCTGCTCCAC 233

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
                |  ||   ||| |  || |||   ||| | ||  |||||| || ||||| ||||||
Db        234 CACGGCAAGCCCATCTCCGAGTCTCTGATCGTCGTCCAGTACATCGACGAGGTCTGGTCC 293

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
              |   || |   || |||| ||     | ||  | |  || ||       || || |||
Db        294 TCCGGCGCGCCGGCCTTCCTCCCCGTCGACGCTCACGCCCGCGCCGTCCAGCGGTTCTGG 353

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
              |    ||| |||| ||||||  || | | |   ||   | | ||   |       ||
Db        354 GCGCAGTACGTCGACGACAAGCTGCCTTGGGC---GATCCGCATACTGAAGGGAACGGAC 410

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
              || | ||  | | |||        | |||   |   |  |  ||  ||    | || ||
Db        411 GACGGGGGCATGGAGCAGGCGGCGGGGCAGCTGTCCGCGGCCCTGCAGCTCCTAGAGGAG 470

Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
              |  |||     ||  |||  | ||||||  ||| ||||| || || |   | || |   |
Db        471 GCTTTCGCGCAGCTCAGCCAGGGGAAGCGCTACTTCGGCGGGGACAGCGTCGGGTACCTG 530

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
              |||||  | | | |        | |  | |   |  |||| ||    |
Db        531 GACATCGCTCTGGTGTCGCATGTCGGCTGGGTGAAGGCGGTGGAGAAGATCGCCGGGGTC 590

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
              |  |||| ||  |  | |||  |||      ||       |||||||   | |   |||
Db        591 ACCCTGCTGGACAAGGCCAAGGTCCCGAACCTGGTGGCGTGGGCTGATCGTCTGTGTGCC 650

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTT 734
              | || || |||   | |   | || ||  | || ||| | ||||||||
Db        651 CACCCGGCCGTGGTCGACGCCATCCCTGACGCGGACAAGTTCGTTGAGTT 700


RESULT 11
US-09-248-335-61
; Sequence 61, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 61
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-61

  Query Match            15.5%;  Score 142.6;  DB 3;  Length 900;
  Best Local Similarity  53.3%;  Pred. No. 2.4e-30;
  Matches  348;  Conservative    0;  Mismatches  299;  Indels    6;  Gaps    2;

Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
              ||| | ||    | | || |  |  | ||    |||   || || || |   |  | ||
Db         28 ATGTCGGAGGCCGCCGTGCGAGTGATCGGCCTATGGCCGAGCCCGTTCGTGATCCGCGTC 87

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
              | ||| || || || || |||  | |||| || ||| | || || |||     ||     |
Db         88 CTGATCGCGCTGAAGCTGAAGCATGTGGAGTACGAGTTCGTGGAGGAGGTGGTGGGCAGC 147

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
              || || ||  |||| ||    ||| ||||| ||||||||||| ||||| ||   |  ||||
Db        148 AAGAGCGAGCTGCTGCTCGCGTCGAACCCGGTGCACAAGAAGATCCCCGTCCTGCTCCAC 207

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
                |  ||   | | |  || ||  |||||| ||| ||||||| || ||||| ||||||
Db        208 CACGGCAAGCCCCTCTCCGAGTCCCTAATCATCGTTCAGTACATCGACGAGGTCTGGTCC 267

Qy        325 AACAATGCTC---TCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTT 381
              |   || |      |||||| ||     | | || |  || |||       || ||
Db        268 TCCGGCGCGCCGGCGGCCATCCTCCCCGCCGACCCTTACGCGCGCGCTGTCCAGCGGTTC 327

Qy        382 TGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAA 441
              ||||     ||| |||| ||||||  ||   || |   ||   | || ||   |
Db        328 TGGGCGCAGTACGTCGACGACAAGATGCACCCGGC---GATCCGCGTACTGAAGGGAACG 384

Qy        442 GATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAA 501
              | || | ||  ||| |||          | |||    |   |  |  ||  ||   ||||
Db        385 TACGACGGGGACAAGGAGCAGGCGGCGGGGCAGCTGTCCGCGGCCCTGCAGCTCCTGGAG 444

Qy        502 GAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTT 561
              || |  |||    ||  | |  | ||||||  ||| ||||| || || |   | || |
Db        445 GAGGCTTTCGCGCAGCTCGGCCAGGGGAAGCGCTACTTCGGCGGGGACAGCGTCGGGTAC 504

Qy        562 GTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAA 621
              | |||||  |    | |          | |  | |   |  |||| ||    |
Db        505 CTGGACATCGCCCTGGTGTCGCACGTCGGCTGGGTGAAGGCGGTGGAGAAGATCGCGGGG 564

Qy        622 AGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCT 681
              |   |||| || ||  |||||   ||      ||       ||||||||     |
Db        565 GTCACTCTGCTGGACGAGGCGAAGGTTCCCAACCTGGTGGCGTGGGCTGACCGGCTGTGC 624

Qy        682 GCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTT 734
              ||  | || || |||  || |    | || ||  | || ||| | ||||||||
Db        625 GCCCACCCGGCCGTGGTGGACGCGATCCCTGACGCCGACAAGTTCGTTGAGTT 677


RESULT 12
US-09-247-373B-53
; Sequence 53, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
;  APPLICANT:  MCGONIGLE, BRIAN
;  APPLICANT:  O'KEEFE, DANIEL
;  TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE:  CL-1108-A
;  CURRENT APPLICATION NUMBER: US/09/247,373B
;  CURRENT FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  08/924,747
;  PRIOR FILING DATE:  1997-09-05
;  NUMBER OF SEQ ID NOS:  56
;  SOFTWARE:  Microsoft Office 97
; SEQ ID NO 53
;   LENGTH: 841
;   TYPE: DNA
;   ORGANISM: SOYBEAN
```

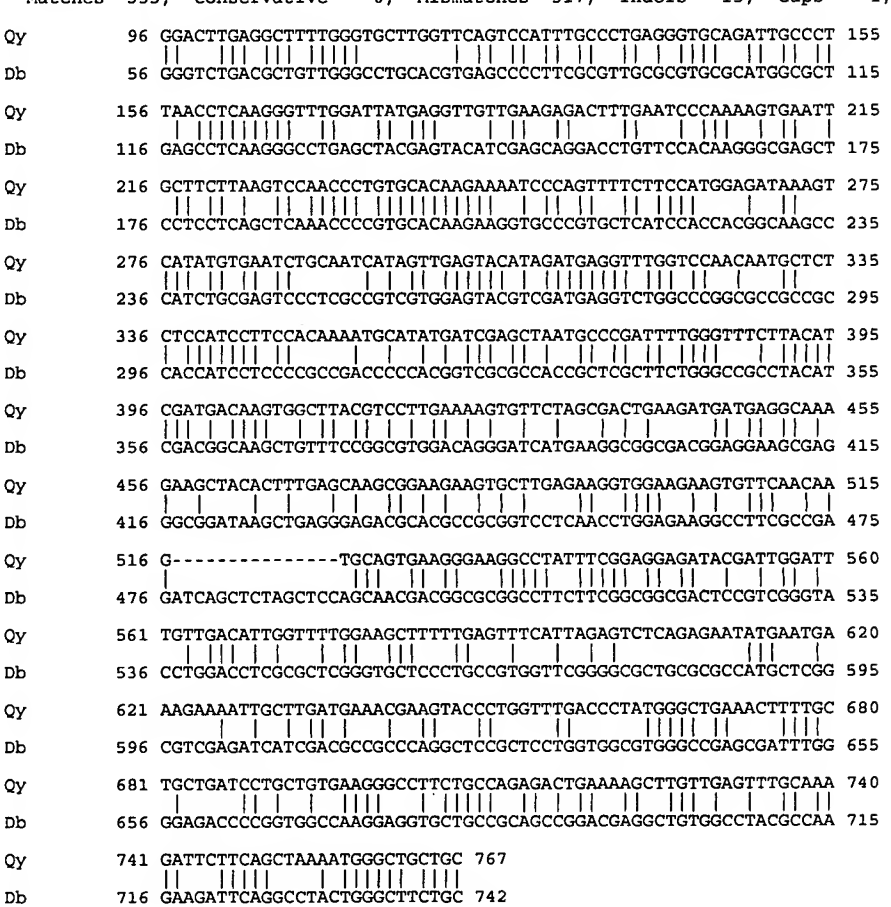
US-09-247-373B-53

Query Match 15.3%; Score 140; DB 3; Length 841;
Best Local Similarity 52.1%; Pred. No. 1.2e-29;
Matches 367; Conservative 0; Mismatches 325; Indels 12; Gaps 2;



RESULT 13
US-09-248-335-41
; Sequence 41, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 41
; LENGTH: 860
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-41

Query Match 15.2%; Score 139.8; DB 3; Length 860;
Best Local Similarity 51.7%; Pred. No. 1.4e-29;
Matches 355; Conservative 0; Mismatches 317; Indels 15; Gaps 1;



RESULT 14
US-09-248-335-71
; Sequence 71, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 71
; LENGTH: 1013

```
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-71

  Query Match             14.3%;  Score 131.4;  DB 3;  Length 1013;
  Best Local Similarity   55.2%;  Pred. No. 3.5e-27;
  Matches  280;  Conservative    0;  Mismatches  221;  Indels    6;  Gaps    1;

Qy         65 ACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCA 124
Db         83 ATCTGGAGAGGGAAACAATGGCGGGCGAGGAGGGTCTTAAGGTCCTCGGCCTGCAGGTGA 142

Qy        125 GTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTG 184
Db        143 GCCCGTTCGTGCTCCGCGTGTGCTTGGCGCTGAACATGAAAGGAGTGAGTTACGAGTACG 202

Qy        185 TTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGA 244
Db        203 TCGAGGAGGACATATCCAACAAGAGTGAGCTCCTGCTCAAGTCCAACCCGGTGCACAAGA 262

Qy        245 AAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGT 304
Db        263 AGGTGCCCGTGCTCATCCACAACGGTAAGCCCATCTGCGAGTCACTCGTCATCATGCAGT 322

Qy        305 ACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATC 364
Db        323 ACGTCGACGAGCTGTTCGCCGGC------CGGCCGATCCTCCCAACCGACCCCTACGAGC 376

Qy        365 GAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAA 424
Db        377 GCGCCACTGCTCGCTTCTGGGCTGCCTACGCCGACGACAAGTTGTTTCCAGCGTGGTACG 436

Qy        425 GTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAG 484
Db        437 GCATGGTGAAGGCCCAGCCGGAGGAGGAGAGGGCGGAGAAGGCGAAGGAGACGCTCGCCG 496

Qy        485 TGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAG 544
Db        497 CCATCGAGCACATGGAAGTGACCTTCGCCAAGTGCTCCGGCGGCAACGCCTTCTTCGGTG 556

Qy        545 GAGATACGATTGGATTTGTTGACATTG 571
Db        557 GCGACTCCATCGGCTACGTCGACATCG 583

RESULT 15
US-09-248-335-63
; Sequence 63, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 63
;   LENGTH: 872
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-63

  Query Match             14.2%;  Score 130.8;  DB 3;  Length 872;
  Best Local Similarity   49.1%;  Pred. No. 4.8e-27;
  Matches  403;  Conservative    0;  Mismatches  412;  Indels    5;  Gaps    2;

Qy        101 TGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACC 160
Db         54 TGAAGCTGATCGGGCAGTACGGGAGCGCGTTCGTGACGAGGGTGAAGCTTGCTCTCAGCC 113

Qy        161 TCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTC 220
Db        114 TCAAGGGGCTGAGCTACGAGTACGTCGAGGAGGATCTCAGAAACAAGAGCGCGCTCCTCC 173

Qy        221 TTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATAT 280
Db        174 TCAGCTCCAACCCGGTGCACAAGGCGGTTCCAGTGCTGATCCACAGAGGCAAGCCTATCT 233

Qy        281 GTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCA 340
Db        234 GCGAGTCGCAGGTCATCGTGCAGTACATCGACGAGGCCTTTGCCGGCATCGGCCCGCCCC 293

Qy        341 TCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATG 400
Db        294 TCCTCCCGGCCGACCCCTACGAACGCTCGGTGGCCCGTTTCTGGGCTGCCTTCATTGAAG 353

Qy        401 ACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGC 460
Db        354 ACAAGCTTGTGTCCCCGTGGGACCGAGTGTTCCGGGCGAAGACGGAGGACGAGAGGGAAG 413

Qy        461 TACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCA 520
Db        414 AGGCGATGAAGCAGATGCTTGCGGCAGTGGACGCTCTGGAGGGAGCACTGAAGGAGGGGA 473

Qy        521 GTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAA 580
Db        474 G---ACCCAGACCCTTCTTCGGCGGCGACAGCGTCGGGTACGTGGACGTCGTTCTGGGCG 530

Qy        581 GCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAA 640
Db        531 GTGCCGTCTCGTACGCCAAGGGGCACGACGCGCTCTTCGGTTCCGAGCTCATCGACGCCG 590

Qy        641 CGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGG 700
Db        591 CCAAGACGCCGCTCCTGGCCGCGTGGATGGAGCGCTTCTGCGAGCTCGACGCGGCCAAGG 650

Qy        701 GCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGG 760
Db        651 CGGTCCTGCAGGACGTCGATAGAGTGGTCCAGTACGGCAAGATGCTGATCGCCAAGAATG 710

Qy        761 CTGCTGCAGCTGCTGCAAAGTAAATGGAATCAAATTAAT--TGCTGGATGAATTTCAAAA 818
Db        711 CTGCTGCCACTCGTCAGGCGTAGTGTTTTTCTGATCGATCAGCTTGTATGTATATGAATT 770

Qy        819 ATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAA 878
Db        771 GAACTTGTAAAACCAAATCGTTCAAGTTTGATGGTAAGTTCCATGTTAGAAAAAAAAAAA 830

Qy        879 GTCAAAATAAATGTTATGATAATATAGTAAAAAAAAAAAA 918
Db        831 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 870

Search completed: October 13, 2004, 06:54:12
Job time : 90 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 12, 2004, 23:53:06 ; Search time 2946 Seconds
(without alignments)
9305.323 Million cell updates/sec

Title: US-10-088-945A-14
Perfect score: 918
Sequence: 1 cataaaactccacatttcct.....aatatagtaaaaaaaaaaa 918

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	592.2	64.5	614	14	CF922626	CF922626 gmrhRww24
	2	486.4	53.0	579	13	BU084645	BU084645 sar26a06.
c	3	475.2	51.8	691	10	BE823032	BE823032 GM700019B
c	4	461.6	50.3	604	14	CD406010	CD406010 Gm_ck3016
	5	461.2	50.2	614	10	AW471665	AW471665 si14c11.y
	6	453.6	49.4	566	12	BG157232	BG157232 sab35d02.
	7	439.4	47.9	601	13	BU764540	BU764540 sas03a07.
	8	438	47.7	581	12	BI427624	BI427624 sah82d09.
c	9	436.8	47.6	458	10	AW349581	AW349581 GM210005A
c	10	422	46.0	707	14	CD400374	CD400374 Gm_ck2222
	11	414.2	45.1	553	13	BU091596	BU091596 st75a10.y
	12	413.6	45.1	534	10	BG043636	BG043636 sv39c03.y
	13	413.4	45.0	601	10	AW760276	AW760276 sl48c04.y
	14	402	43.8	630	10	BE556267	BE556267 sp99e11.y
	15	382.8	41.7	502	12	BG653588	BG653588 sad76c09.
	16	373.8	40.7	627	13	BQ140801	BQ140801 NF043A11P
	17	371.6	40.5	539	10	BG047474	BG047474 saa84b05.
	18	370.4	40.3	372	9	AI440996	AI440996 sa63f01.y
	19	364.8	39.7	616	10	BF520786	BF520786 EST458259
	20	356.4	38.8	489	10	AW734511	AW734511 sk95d10.y
	21	354.4	38.6	583	14	CA953253	CA953253 sav54d11.
	22	352.8	38.4	597	10	BF519757	BF519757 EST457221
	23	341.4	37.2	671	14	CB893410	CB893410 EST646202
	24	329.8	35.9	606	10	BF639128	BF639128 NF080C10P
	25	326.6	35.6	530	14	CA953342	CA953342 sav55e10.
	26	322.8	35.2	500	13	BQ610460	BQ610460 sap42a12.
	27	319.4	34.8	514	13	BQ297110	BQ297110 sao38d01.
c	28	315.8	34.4	775	14	CA923160	CA923160 EST640878
	29	312.8	34.1	741	13	BU861347	BU861347 S001A07 P
c	30	311.4	33.9	458	14	CD396596	CD396596 Gm_ck1687
	31	308.6	33.6	592	10	AW693000	AW693000 NF058C03S
	32	303.8	33.1	504	10	BF716119	BF716119 saa14h02.
	33	297	32.4	841	14	CF510854	CF510854 CAbud0001
	34	294.8	32.1	790	14	CF510677	CF510677 CAbud0001
	35	291.4	31.7	456	10	BF068201	BF068201 st88e11.y
	36	291	31.7	695	14	CF837076	CF837076 UCRCS03_0
	37	289.8	31.6	401	10	BE801758	BE801758 sr30h11.y
c	38	288.8	31.5	721	14	CF837075	CF837075 UCRCS03_0
	39	280.2	30.5	721	14	CA411506	CA411506 E808_F -P
	40	278.4	30.3	639	14	CD278180	CD278180 T143C0183
	41	278.2	30.3	666	14	CK111039	CK111039 Q008B07 P
	42	277.8	30.3	864	14	CF206076	CF206076 RR890915I
	43	275.4	30.0	341	10	BE190293	BE190293 so10b09.y
	44	270.6	29.5	621	14	CF503910	CF503910 USDA-FP_1
	45	267.4	29.1	636	13	BU888339	BU888339 P006G01 P

ALIGNMENTS

RESULT 1
CF922626
LOCUS CF922626 614 bp mRNA linear EST 05-NOV-2003
DEFINITION gmrhRww24-13-T7_D07_1_057 Soybean root hair subtracted cDNA library gmrhRww24 Glycine max cDNA, mRNA sequence.
ACCESSION CF922626
VERSION CF922626.1 GI:38193420
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 614)
AUTHORS Scheffler,B.E., Huang,S., Liu,X., Nguyen,H., Duke,M. and Stacey,G.
TITLE Expressed sequence tags from soybean root hair subtractive cDNA library
JOURNAL Unpublished (2003)
COMMENT Contact: Gary Stacey
University of Missouri
108 Waters Hall, Columbia, MO 65211, USA
Tel: 573-884-4752
Fax: 573-882-0588
Email: staceyg@missouri.edu
Single pass sequence
Seq primer: T7.

```
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /cultivar="Williams 82"
                     /db_xref="taxon:3847"
                     /tissue_type="root hairs"
                     /clone_lib="Soybean root hair subtracted cDNA library
                     gmrhRww24"
                     /note="Organ: root hairs; Vector: pCR2-1 Topo; cDNA clones
                     generated from soybean root hair tissue treated with
                     Bradyrhizobium japonicum for 24 hours."
ORIGIN

  Query Match             64.5%;  Score 592.2;  DB 14;  Length 614;
  Best Local Similarity   99.3%;  Pred. No. 1.1e-105;
  Matches  605;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        307 ATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGA 366
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db          1 ATAGATGAGGTTTGGTCC-ACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGA 59

Qy        367 GCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGT 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGT 119

Qy        427 GTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTG 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 GTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTG 179

Qy        487 CTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGA 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 CTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGA 239

Qy        547 GATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCA 299

Qy        607 GAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGG 666
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        300 GAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGCACCCTGGTTTGACCCTATGG 359

Qy        667 GCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTT 726
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        360 GCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCCGCCAGAGACTGAAAAGCTT 419

Qy        727 GTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAGTAAATG 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAGTAAATG 479

Qy        787 GAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAG 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAG 539

Qy        847 GCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGATAATATAGT 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGATAATATAGT 599

Qy        907 AAAAAAAAA 915
              ||||||| |
Db        600 AAAAAAATA 608


RESULT 2
BU084645
LOCUS       BU084645                579 bp    mRNA    linear   EST 27-AUG-2002
DEFINITION  sar26a06.y1 Gm-c1049 Glycine max cDNA clone SOYBEAN CLONE ID:
            Gm-c1049-7691 5' similar to TR:Q9SEK1 Q9SEK1 GLUTATHIONE
            S-TRANSFERASE 3 ;, mRNA sequence.
ACCESSION   BU084645
VERSION     BU084645.1  GI:22525834
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 579)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact: ccu@resgen.com web site:
            www.resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1. .579
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="SOYBEAN CLONE ID: Gm-c1049-7691"
                     /tissue_type="whole seedlings of greenhouse grown plants"
                     /dev_stage="3 week old"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1049"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The Clark NIL was constructed and seed was provided
                     by Dr. J. Specht, University of Nebraska (Shoemaker and
                     Specht, 1995). The cDNA library was constructed from mRNA
                     isolated from whole seedlings of 3 week old greenhouse
                     grown plants. Complementary DNA was synthesized from mRNA
                     using a primer consisting of a poly(dT) sequence with a
                     XhoI restriction site and a 3' anchor. EcoRI adapters were
                     ligated to the blunt-ended cDNA fragments followed by XhoI
                     digestion. The cDNA fragments were directionally cloned
                     into the EcoRI-XhoI restriction site of the pBluescript
                     vector. The ligated cDNA fragments were transformed into
                     DH10B host cells (GibcoBRL). The library was constructed
                     in cooperation with Dr. Paul Keim's laboratory at Northern
                     Arizona University."
ORIGIN

  Query Match             53.0%;  Score 486.4;  DB 13;  Length 579;
  Best Local Similarity   91.5%;  Pred. No. 5.2e-85;
  Matches  527;  Conservative    0;  Mismatches   46;  Indels    3;  Gaps    1;

Qy         79 ACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTG 138
              |||||||||||| |||||||||||||||||||||||||||||||||||||| ||||
Db          1 ACAGTGATGGCTCAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGTCCTG 60

Qy        139 AGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTG 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTG 120

Qy        199 AATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTC 180

Qy        259 TTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTT 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTT 240
```

```
Qy        319 TGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGA 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGA 300

Qy        379 TTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACT 438
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||| | ||
Db        301 TTTTGGGTTTCTTACATCGATGACAAGTGGGTTACGTCCTTGAAAAGTGTTCTACTGGCT 360

Qy        439 GAAGAT---GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAG 495
              | ||||   || ||||||||||||   |||||||| | ||  |||||||| |||||||
Db        361 GGAGATGACGACGAGGCAAAGAAATCACACTTTGTGGAAATGGAAGAAGCACTTGAGAGG 420

Qy        496 GTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATT 555
              ||||||||||||||||||||||||||||||| ||| | || || ||||| ||||| |||
Db        421 ATGGAAGAAGTGTTCAACAAGTGCAGTGAAGGAAAGACATACTTTGGAGGGGATACAATT 480

Qy        556 GGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATG 615
              ||| |||||||||||| |||||| ||| ||| |||||  || |||||   |||  | |||
Db        481 GGAATTGTTGACATTGTTTTTGGGAGCCTTTGGAGTTGGATGAGAGTGATAGAAGAGATG 540

Qy        616 AATGAAAGAAAATTGCTTGATGAAACGAAGTACCCT 651
              |||| ||||||| || |||||||| | ||| |||||
Db        541 AATGGAAGAAAAGTGTTTGATGAAGCAAAGAACCCT 576
```

```
RESULT 3
BE823032/c
LOCUS       BE823032                691 bp    mRNA    linear   EST 24-MAY-2001
DEFINITION  GM700019B20G8 Gm-r1070 Glycine max cDNA clone Gm-r1070-7624 3',
            mRNA sequence.
ACCESSION   BE823032
VERSION     BE823032.1  GI:10255266
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,
            Erpelding,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.
  TITLE     A Functional Genomics Program for Soybean (NSF 9872565)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: AW471675 corresponding to Gm-c1029-1203 (5')
            Contact: Vodkin, L.O., PI, A Functional Genomics Program for
            Soybean (NSF 9872565)
            Lewin, H. A., Director, Keck Center for Comparative and Functional
            Genomics
            University of Illinois
            Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
            Tel: (217) 244-6147
            Fax: (217) 333-4582
            Email: l-vodkin@uiuc.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134. For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact:clones@genomesystems.com or info@genome
            ystems.com web site:www.genomesystems.com
            Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="Gm-r1070-7624"
                     /clone_lib="Gm-r1070"
                     /note="The library Gm-r1070 is a sequence-driven, reracked
                     set of 9,216  clones selected from cDNA libraries from
                     various tissues and stages  of development of soybean that
                     represent 2,639 sequences from  immature cotyledons, 1,770
                     from immature seed coats, 3,938 from  flowers, and 869
                     from young pods.  The 5' ESTs of the source clones from
                     the different libraries was  used to select singletons, or
                     a representative of each contig, which  were reracked to
                     form library Gm-r1070.  The cDNA clones of the reracked
                     Gm-r1070 library were then sequenced  at the 3' end.  The
                     contig analysis to select unique genes was performed by
                     the  laboratory of Ernest Retzel, Center for Computational
                     Genomics and  Bioinformatics, University of Minnesota,
                     http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
                     .  Reracking was performed by Genome Systems, St. Louis,
                     http://www.genomesystems.com, and 3' sequencing by the
                     Keck Center  for Comparative and Functional Genomics,
                     University of Illinois,
                     http://www.life.uiuc.edu/biotech/keck.html.  Note: The
                     corresponding 5' EST from each clone in the Gm-r1070
                     library is listed in the 'OTHER EST' field.  The detailed
                     information on the source library for each clone can  also
                     be obtained by referring to the Genome Systems clone ID of
                     the  original cDNA library that is also listed under
                     'OTHER EST'."
ORIGIN
```

```
  Query Match             51.8%;  Score 475.2;  DB 10;  Length 691;
  Best Local Similarity   84.2%;  Pred. No. 7.7e-83;
  Matches  567;  Conservative    0;  Mismatches   95;  Indels   11;  Gaps    4;

Qy        234 TGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAAT 293
              |  |  |||||  ||||||||||||||||||   |||||||||| |||| || ||||
Db        691 TNNGNNNAAGAANNTCCCAGTTTTCTTCCATNNNGATAAAGTCATTTGTGNNTCNNCAAT 632

Qy        294 CATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAA 353
              |||||  |||||||| |||||||     |  | |||| ||   ||||  |||||||||
Db        631 CATAGNNGAGTACATTGATGAGGCNNNG---ACTAATGTTCCNNCCATNNNTCCACAAAA 575

Qy        354 TGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTAC 413
              ||| |||||||| |||||||| ||||||||| || | ||||| ||||| |||||| ||||
Db        574 TGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATTGATGAGAAGTGGTTTAC 515

Qy        414 GTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCA 473
              |||||||| |||||||||| | |||||||||||||||||||||||| ||||||||||||
Db        514 GTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAGAAGCCACACTTTGAGCA 455

Qy        474 AGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGC 533
              ||| ||||||| |||||||| | |||||||||||||||||||||| ||||||||||||||
Db        454 AGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAGTACAGTGAAGGGAAGGC 395

Qy        534 CTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTT 593
              |||||||||||||||| |||||||| |||||||||||||| ||||| |||||||
Db        394 CTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTTGGGAGCTTCTTGAGTTG 335

Qy        594 CATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGG 653
              || ||||||  |||| | |||| || |||||||||||||||||||||| |||| |||||||
Db        334 GATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGATGAAAAGAAGCACCCTGG 275

Qy        654 TTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGA 713
              |||||||| |||||||||||| |||||||||||||||||||||||||||| |||| |||||
Db        274 TTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCCTGCTGTGAAGGGCATTCTTCCAGA 215

Qy        714 GACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCA----- 768
              |||||| |||||||||||||||| ||||||||||||||||||||| ||||||||
Db        214 GACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCAGCTAAAATGGACTGCTGCAGCAGC 155

Qy        769 -GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTG 827
               || |||||||||||||||||||||||||||||||  || | ||||  ||||||||||
Db        154 TGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCGAGA-GTATTTTCAAAATTGTTGTC 96

Qy        828 CAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATA 887
              |||||| ||| ||||| |||||||| |||||||||||||| |||||||||| |   ||
Db         95 CAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAACTTTATTTATTTAAAAGTTATTTTA 37

Qy        888 AATGTTATGATAA 900
              |     |||  ||
Db         36 ATNNNAATGTAAA 24
```

```
RESULT 4
CD406010/c
LOCUS       CD406010                 604 bp    mRNA    linear   EST 07-JUN-2003
DEFINITION  Gm_ck30160 Soybean induced by Salicylic Acid Glycine max cDNA 3',
            mRNA sequence.
ACCESSION   CD406010
VERSION     CD406010.1  GI:31463982
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Tian,A.-G., Wang,J., Cui,P., Han,Y.-J., Xu,H., Cong,L.-J.,
            Huang,X.-G., Wang,X.-L., Jiao,Y.-Z., Wang,B.-J., Wang,Y.-J.,
            Zhang,J.-S., Chen,S.-Y. and Yu,J.
  TITLE     Soybean Expressed Sequence Tags Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Chen S-Y
            Plant Biotechnology Laboratory
            Institute of Genetics and Developmental Biology, CAS, China
            Datun road, Beijing 100101, China
            Tel: 86-10-64886859
            Fax: 86-10-64873428
            Email: sychen@genetics.ac.cn
            Email: sychen@genetics.ac.cn
            Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .604
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /cultivar="Kefeng 1"
                     /db_xref="taxon:3847"
                     /tissue_type="Seedlings"
                     /dev_stage="two-week seedlings"
                     /lab_host="XL1-Blue MRF' strain"
                     /clone_lib="Soybean induced by Salicylic Acid"
                     /note="Vector: pBluescript SK+; Site_1: EcoR I; Site_2:
                     Xho I; The cDNA library was constructed by He, C-Y from
                     mRNA isolated from two-week seedlings (cultivar Kefeng 1)
                     treated by spraying 2.0mM salicylic acid for 24, 36, 48
                     and 72 h. Complementary DNA was synthesized from mRNA
                     using a primer consisting of a poly(dT) sequence with a
                     XhoI striction site. EcoRI adapters were ligated to the
                     blunt-ended cDNA fragments followed by XhoI digestion. The
                     cDNA fragments were directionally cloned into the
                     EcoRI-XhoI restriction site of the pBluescript vector. The
                     ligated cDNA fragments were transformed into XL1-Blue MRF'
                     host cells (Stratagene)."
ORIGIN

  Query Match           50.3%;  Score 461.6;  DB 14;  Length 604;
  Best Local Similarity  89.1%;  Pred. No. 3.6e-80;
  Matches  545;  Conservative    0;  Mismatches   59;  Indels    8;  Gaps    4;

Qy        260 TCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTT 319
              |||||||||||||||||| |||||||||||||||||||||||||||||| ||||||| ||
Db        604 TCCATGGAGATAAAGTCATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTT 545

Qy        320 GGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGAT 379
              ||   |  |||| || |||||||||||||||||||| |||||||| |||||||| ||||
Db        544 GG---ACTAATGTTCCCTCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGAT 488

Qy        380 TTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTG 439
              ||||| || | ||||| ||||| |||||| |||||||||||| |||||||||| | |||
Db        487 TTTGGTTTGCCTACATTGATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTG 428

Qy        440 AAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGG 499
              |||||||||||||||||||| |||||||||||||| ||||||| |||||||| | |||

Db        427 AAGATGATGAGGCAAAGAAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGG 368

Qy        500 AAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGAT 559
              |||||||||||||||||| |||||||||||||||||||||||||||||||  |||||||
Db        367 AAGAAGTGTTCAACAAGTACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGAT 308

Qy        560 TTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATG 619
              |  |||||||||||||||| ||||| |||||||  || ||||||  |||| | |||| ||
Db        307 TCATTGACATTGGTTTTGGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTG 248

Qy        620 AAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTG 679
              |||||||||||||||||||| |||| |||||||||||||| |||||||||||| ||||
Db        247 GAAGAAAATTGCTTGATGAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTG 188

Qy        680 CTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAA 739
              |||||||||||||||||||||| |||| |||||||||| |||||||||||||||||| |
Db        187 CTGCTGATCCTGCTGTGAAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCA 128

Qy        740 AGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAGTAAATGGAATCAAATTAAT 799
              |||||||||||||||||||| ||||||||  ||||||||||||||||||||||||||||
Db        127 AGATTCTTCAGCTAAAATGGACTGCTGCA---GCTGCAAAGTAAATGGAATCAAATTAAT 71

Qy        800 TGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTG 859
              |||  || | ||||  |||||||||||| |||||| ||| ||||| |||||||| ||||||
Db         70 TGCGAGA-GTATTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTG 13

Qy        860 CAACTTTATATA 871
              |||||||| ||
Db         12 CAACTTTATTTA 1


RESULT 5
AW471665
LOCUS       AW471665                 614 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  si14c11.y1 Gm-c1029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1029-1197 5' similar to TR:O04941 O04941 GLUTATHIONE
            S-TRANSFERASE TSI-1 ;, mRNA sequence.
ACCESSION   AW471665
VERSION     AW471665.1  GI:7041771
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 614)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            High quality sequence stop: 441.
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1029-1197"
                     /tissue_type="very young cotyledons of greenhouse grown
```

plants"
/lab_host="DH10B"
/clone_lib="Gm-c1029"
/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; This cDNA library was constructed from mRNA isolated from very young cotyledons (20-50mgs fresh weight) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

ORIGIN

Query Match 50.2%; Score 461.2; DB 10; Length 614;
Best Local Similarity 89.3%; Pred. No. 4.3e-80;
Matches 509; Conservative 0; Mismatches 58; Indels 3; Gaps 1;

```
Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| |||||||||  |  ||||||| ||||||||||||||||||||||||||||||
Db         48 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 107

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || |||||||||||||||||||| ||||||||||||||||  ||
Db        108 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 167

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 227

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        228 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 284

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||| |||||||| |||||||| ||||||||| || | |||||
Db        285 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 344

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        345 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 404

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        405 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 464

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | |||||||||||| |||||||||||||||||||  |||||||  ||||||||||||||
Db        465 TACAGTGAAGGGAATGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 524

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||  || ||||||  |||| | |||| || |||| ||||||||||||
Db        525 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGATAATTGCTTGAT 584

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGG 666
              || | |||| ||||||| ||||||| ||||
Db        585 GATAAGAAGCACCCTGGGTTGACCCAATGG 614
```

RESULT 6
BG157232
LOCUS BG157232 566 bp mRNA linear EST 28-NOV-2001
DEFINITION sab35d02.y1 Gm-c1026 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1026-3604 5' similar to TR:Q9SEK1 Q9SEK1 GLUTATHIONE S-TRANSFERASE 3 ;, mRNA sequence.
ACCESSION BG157232
VERSION BG157232.1 GI:12690896
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 566)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 484.
FEATURES Location/Qualifiers
source 1. .566
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1026-3604"
/tissue_type="Senescing leaves, mature plants, greenhouse grown."
/lab_host="DH10B"
/clone_lib="Gm-c1026"
/note="Vector: pT7T3-Pac (Pharmacia); Site_1: EcoRI; Site_2: HindIII; This cDNA library was constructed from mRNA isolated from senecsing leave tissue of mature greenhouse grown plants. Complementary DNA was synthesized from mRNA using a 3' anchored poly(dT) primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and HindIII. The cDNA fragments were directionally cloned into the EcoRI-HindIII restriction site of the pT7T3-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed R. Shoemaker and J. Erpelding."

ORIGIN

Query Match 49.4%; Score 453.6; DB 12; Length 566;
Best Local Similarity 89.0%; Pred. No. 1.3e-78;
Matches 502; Conservative 0; Mismatches 59; Indels 3; Gaps 1;

```
Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| |||||||||  |  ||||||| ||||||||||||||||||||||||||||||
Db          6 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 65

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || |||||||||||||||||||| ||||||||||||||||  ||
Db         66 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGACCTG 125

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 185

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        186 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 242

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||| |||||||| |||||||| ||||||||| || | |||||
Db        243 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 302
```

```
Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        303 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 362

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        363 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 422

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | ||||||||||||||||||||||||||||||  ||||||||  |||||||||| |||
Db        423 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGGTTT 482

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
              || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||||||||
Db        483 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTGCTTGAT 542

Qy        637 GAAACGAAGTACCCTGGTTTGACC 660
                ||  ||| ||||| ||||| ||
Db        543 TGAAAAAAGCACCCTTGTTTGGCC 566
```

RESULT 7
BU764540

LOCUS	BU764540 601 bp mRNA linear EST 10-OCT-2002
DEFINITION	sas03a07.y2 Gm-c1080 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1080-2701 5' similar to TR:Q9SEK1 Q9SEK1 GLUTATHIONE S-TRANSFERASE 3 ;, mRNA sequence.
ACCESSION	BU764540
VERSION	BU764540.1 GI:23732736
KEYWORDS	EST.
SOURCE	Glycine max (soybean)
ORGANISM	Glycine max Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE	1 (bases 1 to 601)
AUTHORS	Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com Seq primer: -40RP from Gibco High quality sequence stop: 455.

FEATURES Location/Qualifiers

```
     source          1. .601
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="SOYBEAN CLONE ID: Gm-c1080-2701"
                     /tissue_type="Roots of 8 day old 'Bragg' supernodulating
                     mutant NTS382 seedlings"
                     /dev_stage="8 days old"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1080"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The mRNA was isolated from roots of 8 day old
                     'Bragg' supernodulating mutant NTS382 seedlings that were
                     infected with Bradyhizobium japonicum, strain USDA 110, 72
                     hours prior to harvest. Dr. Gary Stacey generously donated
                     the tissue. The roots were flash-frozen in liquid
                     nitrogen. Stratagene's cDNA Synthesis Kit (catalog number
                     200401) was used to synthesize the cDNA. First-strand
                     synthesis was performed with 5-methyl dCTP, hence the
                     ligated cDNA was hemimethylated. A modification of
                     Stratagene's first-strand synthesis primer was used. An
                     'anchor' nucleotide (V=A, C, or G) was added to the 3' end
                     of the primer [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to
                     anchor the primer at the 5' end of the poly(A) tract.
                     After second-strand synthesis, the cDNA ends were filled
                     in with cloned Pfu DNA, ligated to EcoRI adapters and
                     subsequently phosphorylated. The cDNA was then
                     precipitated and redissolved in sterile, RNase-,
                     DNase-free water. The XhoI site within the first- strand
                     synthesis primer was then restricted by digestion with
                     XhoI from Promega (40U/ul); all XhoI sites in the cDNA
                     would be protected by their hemimethylated status. The
                     cDNA constructs were size-fractionated with a 500bp
                     cutoff, using Sephacryl S-500 High Resolution (Pharmacia
                     Biotech) in a 2-mm diameter column and a bed volume of
                     approximately 1ml. The column eluent was precipitated,
                     redissolved, and ligated into Stratagene's pBluescript II
                     XR Predigested vector (pBluescript II SK(+) vector that
                     has been digested with EcoRI and XhoI, and phosphorylated
                     by Stratagene). This library was constucted in the
                     laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at
                     Northern Arizona University."
```

ORIGIN

```
  Query Match             47.9%;  Score 439.4;  DB 13;  Length 601;
  Best Local Similarity   85.2%;  Pred. No. 7.7e-76;
  Matches  507;  Conservative    0;  Mismatches   76;  Indels   12;  Gaps    1;

Qy         75 GTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGC 134
              || || ||  |||||| |||  || ||||||||||||||||||||||||||||| |||
Db          1 GTGAATAGAAATGGCTAAAAACGAGTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATATGC 60

Qy        135 CCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGAC 194
              |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db         61 CCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTCTAGATTATGAGGTTGTTGAAGAGAC 120

Qy        195 TTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGT 254
              |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        121 CTTGAATCCCAAAAGTGATCTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGT 180

Qy        255 TTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGA 314
              | |  |||||||| || || |||||||||||||||||||||||| ||||||||||||||
Db        181 TCTACTCCATGGATATTAATTCATATGTGAATCTGCAATCATAGTAGAGTACATAGATGA 240

Qy        315 GGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGC 374
              |||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db        241 GGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCTTATGACCGAGCTAATGC 300

Qy        375 CCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGC 434
              |||||||||||||||||||||||||||||||  ||||||| | |  | |  | |
Db        301 CCGATTTTGGGTTTCTTACATCGATGACAAGTGGTATACGTCCATAAGGAATACTATTTT 360

Qy        435 GACTG----------AAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGA 482
              | |||           |||||||||||||||||||| |||||||| |  |  ||||||
Db        361 GGCTGAAGCGGCTGATCAAGATGATGAGGCAAAGAAGCCACACTTTGTGAGAATGGAAGA 420

Qy        483 AGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGG 542
              ||  |||||||| | ||||||||||||||||||||||||||||||| | ||| ||||| ||
Db        421 AGCTCTTGAGAGGATGGAAGAAGTGTTCAACAAGTGCAGTGAAGGAAGGGCTTATTTTGG 480

Qy        543 AGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGT 602
              |||||||| |||||| || |||||||| |||||| ||| ||| | |||   | |||||
Db        481 AGGAGATACAATTGGAATTATTGACATTGCTTTTGGGAGCCTTTGGGGTTGGGTGAGAGT 540

Qy        603 CTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTG 657
                |||| | ||||||| ||||||| || |||||||| | ||  |||||| ||||
```

```
Db        541 GATAGAGGAGATGAATGGAAGAAAAGTGTTTGATGAAGCAAAAAACCCTGCTTTG 595


RESULT 8
BI427624
LOCUS       BI427624                581 bp    mRNA    linear   EST 29-NOV-2001
DEFINITION  sah82d09.y1 Gm-c1050 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-c1050-2489 5' similar to TR:Q9SEK1 Q9SEK1 GLUTATHIONE
            S-TRANSFERASE 3 ;, mRNA sequence.
ACCESSION   BI427624
VERSION     BI427624.1  GI:15204856
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            High quality sequence stop: 432.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1050-2489"
                     /tissue_type="leaf tissue at various developmental stages
                     of greenhouse grown plants"
                     /dev_stage="3 week old"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1050"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The Clark NIL was constructed and seed was provided
                     by Dr. J. Specht, University of Nebraska (Shoemaker and
                     Specht, 1995). The cDNA library was constructed from mRNA
                     isolated from leaf tissue at various developmental stages
                     of 3 week old greenhouse grown plants. Complementary DNA
                     was synthesized from mRNA using a primer consisting of a
                     poly(dT) sequence with a XhoI restriction site and a 3'
                     anchor. EcoRI adapters were ligated to the blunt-ended
                     cDNA fragments followed by XhoI digestion. The cDNA
                     fragments were directionally cloned into the EcoRI-XhoI
                     restriction site of the pBluescript vector. The ligated
                     cDNA fragments were transformed into DH10B host cells
                     (GibcoBRL). The library was constructed in cooperation
                     with Dr. Paul Keim's laboratory at Northern Arizona
                     University."
ORIGIN

  Query Match           47.7%;  Score 438;  DB 12;  Length 581;
  Best Local Similarity  90.1%;  Pred. No. 1.5e-75;
  Matches  481;  Conservative    0;  Mismatches   50;  Indels    3;  Gaps    1;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
              ||||||| ||||||||| |  |||||||| ||||||||||||||||||||||||||||||

Db         51 GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGATTGCCCTT 110

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
              |||||||||||| | || ||||||||||||||||||||||||||||||||||||  ||
Db        111 AACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGACCTG 170

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 230

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
              || |||||||||||||||||||||||||||||||| ||||||| ||||   |  |||| || |
Db        231 ATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAATGTTCCC 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
              |||||||||||||||||| |||||||| |||||||| ||||||||| || | |||||
Db        288 TCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGCCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
              ||||| |||||| |||||||||||| |||||||||| | ||||||||||||||||||||
Db        348 GATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGAGGCAAAG 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
              |||| |||||||||||||| ||||||| |||||||| | ||||||||||||||||||||
Db        408 AAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTTCAACAAG 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
              | ||||||||||||||||||||||||||||||  ||||||||  ||||||||||||||
Db        468 TACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACATTGGTTTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTG 630
              || ||||| |||||||  || ||||||  |||| | |||| || ||||||||||
Db        528 GGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATTG 581


RESULT 9
AW349581/c
LOCUS       AW349581                458 bp    mRNA    linear   EST 04-OCT-2000
DEFINITION  GM210005A21F7 Gm-r1021 Glycine max cDNA clone Gm-r1021-1790 3',
            mRNA sequence.
ACCESSION   AW349581
VERSION     AW349581.1  GI:6847291
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,
            Erpelding,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.
  TITLE     A Functional Genomics Program for Soybean (NSF 9872565)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: AI440996
            Contact: Vodkin, L.O., PI, A Functional Genomics Program for
            Soybean (NSF 9872565)
            Lewin, H. A., Director, Keck Center for Comparative and Functional
            Genomics
            University of Illinois
            Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
            Tel: (217) 244-6147
            Fax: (217) 333-4582
            Email: l-vodkin@uiuc.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134. For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888)919-3324 or (314)
            427-3324 or contact:clones@genomesystems.com or info@genome
            ystems.com web site:www.genomesystems.com
            Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Glycine max"
```

/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="Gm-r1021-1790"
/tissue_type="root"
/lab_host="XL10-Gold"
/clone_lib="Gm-r1021"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; Library Gm-r1021 is a sequence-driven, reracked set of the original library Gm-c1004 which was prepared from root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. The Gm-c1004 library was constructed by Dr. Paul Keim & Virginia H. Coryell, Department of Biology, Box5640, Northern Arizona University, Flagstaff, AZ 86011, email: paul.keim@nau.edu, virginia.coryell@nau.edu. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Computational Biology Centers, University of Minnesota, http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html . Reracking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html."

ORIGIN

Query Match 47.6%; Score 436.8; DB 10; Length 458;
Best Local Similarity 97.4%; Pred. No. 2.6e-75;
Matches 444; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy        418 TTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCG 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 TTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCG 397

Qy        478 GAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTAT 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 GAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTAT 337

Qy        538 TTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATT 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 TTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATT 277

Qy        598 AGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTG 657
              |||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db        276 AGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAACCGAAGCACCCTGGTTTG 217

Qy        658 ACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACT 717
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        216 CCCCTATGGGCTGAAACTTTTGCTGATGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACT 157

Qy        718 GAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCA 777
              ||||||| ||||||||| |||||||||| ||| ||||||||||||||||||||||||||
Db        156 GAAAAGCGTGTTGAGTTAGCAAAGATTCTACAGTTAAAATGGGCTGCTGCAGCTGCTGCA 97

Qy        778 AAGTAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTT 837
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 AAGTAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTT 37

Qy        838 ATATCTGAGGCTATGTTTGTTGCAACTTTATATATT 873
              ||||| ||||| |||| ||||||| ||||||||||
Db         36 ATATCAGAGGCAATGTCTGTTGCAGCTTTATATATT 1
```

RESULT 10
CD400374/c
LOCUS CD400374 707 bp mRNA linear EST 07-JUN-2003
DEFINITION Gm_ck22228 Soybean induced by Salicylic Acid Glycine max cDNA 3', mRNA sequence.
ACCESSION CD400374
VERSION CD400374.1 GI:31458346
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 707)
AUTHORS Tian,A.-G., Wang,J., Cui,P., Han,Y.-J., Xu,H., Cong,L.-J., Huang,X.-G., Wang,X.-L., Jiao,Y.-Z., Wang,B.-J., Wang,Y.-J., Zhang,J.-S., Chen,S.-Y. and Yu,J.
TITLE Soybean Expressed Sequence Tags Sequencing
JOURNAL Unpublished (2003)
COMMENT Contact: Chen S-Y
Plant Biotechnology Laboratory
Institute of Genetics and Developmental Biology, CAS, China
Datun road, Beijing 100101, China
Tel: 86-10-64886859
Fax: 86-10-64873428
Email: sychen@genetics.ac.cn
Email: sychen@genetics.ac.cn
Seq primer: T7 primer.
FEATURES Location/Qualifiers
source 1. .707
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Kefeng 1"
/db_xref="taxon:3847"
/tissue_type="Seedlings"
/dev_stage="two-week seedlings"
/lab_host="XL1-Blue MRF' strain"
/clone_lib="Soybean induced by Salicylic Acid"
/note="Vector: pBluescript SK+; Site_1: EcoR I; Site_2: Xho I; The cDNA library was constructed by He, C-Y from mRNA isolated from two-week seedlings (cultivar Kefeng 1) treated by spraying 2.0mM salicylic acid for 24, 36, 48 and 72 h. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI striction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL1-Blue MRF' host cells (Stratagene)."

ORIGIN

Query Match 46.0%; Score 422; DB 14; Length 707;
Best Local Similarity 79.8%; Pred. No. 1.9e-72;
Matches 544; Conservative 0; Mismatches 115; Indels 23; Gaps 3;

```
Qy        209 GTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAG 268
              ||||  |||||||||||||||||||||||||||||||||||||| |  |||||||||
Db        707 GTGATCTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTCTACTCCATGGAG 648

Qy        269 ATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACA 328
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ATAAAGTCATATGTGAATCTGCAATCATAGTAGAGTACATAGATGAGGTTTGGTCCAACA 588

Qy        329 ATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTT 388
              |||||||||||||||||||||| |||||| ||||| ||||||||||||||||||||||||
Db        587 ATGCTCTCTCCATCCTTCCACTAAATGCTTATGACCGAGCTAATGCCCGATTTTGGGTTT 528

Qy        389 CTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTG--------- 439
              |||||||||||||||||||  ||||||| | |  | |  | |   | |||
Db        527 CTTACATCGATGACAAGTGGTATACGTCCATAAGGAATACTATTTTGGCTGAAGCGGCTG 468

Qy        440 ---AAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGG 496
                  |||||||||||||||||||| ||||||| |  |  |||||||  |||||| |
Db        467 ATCAAGATGATGAGGCAAAGAAGCCACACTTTGTGGGAATGGAAGAAGCTCTTGAGAGGA 408

Qy        497 TGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTG 556
              |||||||||||||||||||||||||||||| | ||| ||||| |||||||||| ||||
Db        407 TGGAAGAAGTGTTCAACAAGTGCAGTGAAGGAAGGGCTTATTTTGGAGGAGATACAATTG 348
```

```
Qy        557 GATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGA 616
              || || |||||||||| |||||| ||| ||| | |||   | |||||   |||| | ||||
Db        347 GAATTATTGACATTGCTTTTGGGAGCCTTTGGGGTTGGGTGAGAGTGATAGAGGAGATGA 288

Qy        617 ATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTT 676
              ||| |||||||| || |||||||| | ||  |||||| |||| |   ||||||||| |  |
Db        287 ATGGAAGAAAAGTGTTTGATGAAGCAAAAAACCCTGCTTTGGCTAAATGGGCTGATAAGT 228

Qy        677 TTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTG 736
              || |||||||||||||||||||||  |||| |||||||| |||||||| |||||||||
Db        227 TTTCTGCTGATCCTGCTGTGAAGGGTGTTCTTCCAGAGACTCAAAAGCTTATTGAGTTTG 168

Qy        737 CAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAAAGTAAATGGAATCAAATT 796
              | || |||| | ||| ||||||||||||||||   ||||||||||||||||||| || |
Db        167 CCAAAATTCGTAAGCAAAAATGGGCTGCTGCA---GCTGCAAAGTAAATGGAATTAATTG 111

Qy        797 AATTGCTGGATG--------AATTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGC 848
                  |   ||         | | | ||   ||  | | ||  || |||   |   || |
Db        110 CTAGATTACTTGGGCTTGAAATTGTGAAGTTTTACTTTTCATTTTTTTTCCCTTGTAGAC 51

Qy        849 TATGTTTGTTGCAACTTTATAT 870
              ||||||||||  ||||||| |
Db         50 TATGTTTGTTCTAACTTTATGT 29
```

RESULT 11
BU091596
LOCUS BU091596 553 bp mRNA linear EST 29-AUG-2002
DEFINITION st75a10.y1 Gm-c1054 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1054-43 5' similar to TR:Q9XIF8 Q9XIF8 F23H11.1 PROTEIN. ;, mRNA sequence.
ACCESSION BU091596
VERSION BU091596.1 GI:22541753
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 553)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 439.

FEATURES Location/Qualifiers
source 1. .553
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1054-43"
/tissue_type="Leaf, 3 week old, greenhouse grown"
/lab_host="DH10B"
/clone_lib="Gm-c1054"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from leaf tissue at various developmental stages of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

ORIGIN

Query Match 45.1%; Score 414.2; DB 13; Length 553;
Best Local Similarity 86.0%; Pred. No. 6.5e-71;
Matches 474; Conservative 0; Mismatches 65; Indels 12; Gaps 1;

```
Qy         72 ATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATT 131
              ||||| || ||  |||||| |||  || ||||||||||||||||||||||||||||||||
Db          1 ATAGTGAATAGAAATGGCTAAAAACGAGTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATA 60

Qy        132 TGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGA 191
              |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
Db         61 TGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTCTAGATTATGAGGTTGTTGAAGA 120

Qy        192 GACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCC 251
              ||| ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        121 GACCTTGAATCCCAAAAGTGATCTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCC 180

Qy        252 AGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGA 311
              |||| |  ||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        181 AGTTCTACTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTAGAGTACATAGA 240

Qy        312 TGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAA 371
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||
Db        241 TGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCTTATGACCGAGCTAA 300

Qy        372 TGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCT 431
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||| | |  | | | |
Db        301 TGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGTATACGTCCATAAGGAATACTAT 360

Qy        432 AGCGACTG------------AAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGA 479
                  | |||            |||||||||||||||||||| |||||||| |  |  |||
Db        361 TTTGGCTGAAGCGGCTGATCAAGATGATGAGGCAAAGAAGCCACACTTTGTGAGAATGGA 420

Qy        480 AGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTT 539
              |||||  ||||||| | |||||||||||||||||||||||||||||| | ||| |||||
Db        421 AGAAGCTCTTGAGAGGATGGAAGAAGTGTTCAACAAGTGCAGTGAAGGAAGGGCTTATTT 480

Qy        540 CGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAG 599
              |||||||||| |||||| || |||||||||| || ||| ||| ||| | | ||  | ||
Db        481 TGGAGGAGATACAATTGGAATTATTGACATTGCTTNTGGGAGCCTTTGGGGGTTGGTGAG 540

Qy        600 AGTCTCAGAGA 610
              |||   ||||
Db        541 AGTGATNGAGA 551
```

RESULT 12
BG043636
LOCUS BG043636 534 bp mRNA linear EST 06-DEC-2001
DEFINITION sv39c03.y1 Gm-c1057 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1057-2021 5' similar to TR:Q9XIF8 Q9XIF8 F23H11.1 PROTEIN. ;, mRNA sequence.
ACCESSION BG043636
VERSION BG043636.1 GI:12490115
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Insert Length: 994   Std Error: 0.00
            High quality sequence stop: 424.
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1057-2021"
                     /tissue_type="Degenerating cotyledons, 2 week old
                     seedling"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1057"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from mRNA isolated
                     from degenerating cotyledons of 2 week old seedlings from
                     PI468916. Complementary DNA was  synthesized from mRNA
                     using a primer consisting of a poly(dT) sequence with  a
                     XhoI restriction site. EcoRI adapters were ligated to the
                     blunt-ended cDNA  fragments followed by XhoI digestion.
                     The cDNA fragments were directionally  cloned into the
                     EcoRI-XhoI restriction site of the pBluescript vector. The
                     ligated cDNA fragments were transformed into DH10B host
                     cells (GibcoBRL).  This library was constructed in the
                     laboratory of Dr. Randy Shoemaker."
ORIGIN

  Query Match             45.1%;  Score 413.6;  DB 10;  Length 534;
  Best Local Similarity   87.6%;  Pred. No. 8.6e-71;
  Matches  468;  Conservative    0;  Mismatches   54;  Indels   12;  Gaps    1;

Qy         67 CTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGT 126
              ||| |||||| || ||  |||||| |||  || ||||||||||||||||||||||||||
Db          1 CTGATATAGTGAATAGAAATGGCTAAAAACGAGTTGAGGCTTTTGGGTGCTTGGTTCAGT 60

Qy        127 CCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTT 186
              |||| |||||||||||||||||||||||||||||||||||| | ||||||||||||||||
Db         61 CCATATGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTCTAGATTATGAGGTTGTT 120

Qy        187 GAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAA 246
              |||||||| |||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        121 GAAGAGACCTTGAATCCCAAAAGTGATCTGCTTCTTAAGTCCAACCCTGTGCACAAGAAA 180

Qy        247 ATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTAC 306
              ||||||||| |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCCAGTTCTACTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTAC 240

Qy        307 ATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGA 366
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||
Db        241 ATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCTTATGACCGA 300

Qy        367 GCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGT 426
              |||||||||||||||||||||||||||||||||||||||||||||||  ||||||| | |  | |
Db        301 GCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGTATACGTCCATAAGGAAT 360

Qy        427 GTTCTAGCGACTG------------AAGATGATGAGGCAAAGAAGCTACACTTTGAGCAA 474
                | |   | |||            |||||||||||||||||||| |||||||| |  |
Db        361 ACTATTTTGGCTGAAGCGGCTGATCAAGATGATGAGGCAAAGAAGCCACACTTTGTGGGA 420

Qy        475 GCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCC 534
              ||||||||  |||||||  | ||||||||||||||||||||||||||||||| | |||
Db        421 ATGGAAGAAGCTCTTGAGAGGATGGAAGAAGTGTTCAACAAGTGCAGTGAAGGAAGGGCT 480

Qy        535 TATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTG 588
              ||||| |||||||||||| |||||| || |||||||| |||||| ||| |||||
Db        481 TATTTTGGAGGAGATACAATTGGAATTATTGACATTGCTTTTGGGAGCCTTTTG 534


RESULT 13
AW760276
LOCUS       AW760276                601 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  sl48c04.y1 Gm-c1027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1027-4543 5' similar to TR:O04562 O04562 T7N9.15. ;, mRNA
            sequence.
ACCESSION   AW760276
VERSION     AW760276.1  GI:7692162
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Insert Length: 1392   Std Error: 0.00
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1027-4543"
                     /tissue_type="cotyledons of 3- and 7-day-old Williams
                     seedlings"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1027"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from cotyledons of 3- and 7-day-old Williams seedlings
                     which were propagated on paper towels with distilled
                     water.  The cotyledons were flash-frozen in liquid
                     nitrogen, then lyophilized for 72 hours.  Unequal amounts
                     of mRNA was used for cDNA synthesis.  Stratagene's cDNA
                     Synthexix Kit (catalog number 200401) was used to
                     synthesize the cDNA.  First- stranded synthesis was
                     performed with 5-methyl dCTP, hence the ligated cDNA was
```

hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An anchor nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18] to anchor the primer at the 5' end of the poly(A) tract. After second- strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript(tm) II XR Predigested vector (pBluescript II SK(+) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

ORIGIN

Query Match 45.0%; Score 413.4; DB 10; Length 601;
Best Local Similarity 87.2%; Pred. No. 9.3e-71;
Matches 469; Conservative 0; Mismatches 57; Indels 12; Gaps 1;

```
Qy          61 TTTGACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGG 120
               ||||| ||| |||||| || ||  |||||| |||  || ||||||||||||||||||||
Db          59 TTTGATCTGATATAGTGAATAGAAATGGCTAAAAACGAGTTGAGGCTTTTGGGTGCTTGG 118

Qy         121 TTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 180
               |||||||||| |||||||||||||||||||||||||||||||||||| | |||||||||
Db         119 TTCAGTCCATATGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTCTAGATTATGAG 178

Qy         181 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240
               |||||||||||||| ||||||||||||||||  ||||||||||||||||||||||||||
Db         179 GTTGTTGAAGAGACCTTGAATCCCAAAAGTGATCTGCTTCTTAAGTCCAACCCTGTGCAC 238

Qy         241 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 300
               |||||||||||||| |  ||||||||||||||||||||||||||||||||||||||||
Db         239 AAGAAAATCCCAGTTCTACTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTA 298

Qy         301 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 360
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         299 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCTTAT 358

Qy         361 GATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 420
               || |||||||||||||||||||||||||||||||||||||||||||||  ||||||| |
Db         359 GACCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGTATACGTCCATA 418

Qy         421 AAAAGTGTTCTAGCGACTG----------AAGATGATGAGGCAAAGAAGCTACACTTT 468
               | | |  | |   | |||           |||||||||||||||||||| ||||| |
Db         419 AGGAATACTATTTTGGCTGAAGCGGCTGATCAAGATGATGAGGCAAAGAAGCCACACTNT 478

Qy         469 GAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGG 528
               | |  |  ||||||||  |||||||| | ||||||||||||| ||||||||||||||||||
Db         479 GTGAGAATGGAAGAAGCTCTTGAGAGGATGGAAGAAGTGTACAACAAGTGCAGTGAAGGA 538

Qy         529 AAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTT 586
               ||||| ||||| |||||||||||| |||||| || |||||||| |||||| ||| |||
Db         539 AAGGCTTATTTTGGAGGAGATACAATTGGAATTATTGACATTGCTTTTGGGAGCCTTT 596
```

RESULT 14
BE556267
LOCUS BE556267 630 bp mRNA linear EST 04-DEC-2001
DEFINITION sp99e11.y1 Gm-c1045 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1045-1797 5' similar to TR:O04562 O04562 T7N9.15. ;, mRNA sequence.
ACCESSION BE556267
VERSION BE556267.1 GI:9820757
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 630)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 686 Std Error: 0.00
High quality sequence stop: 415.
FEATURES Location/Qualifiers
source 1. .630
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1045-1797"
/tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
/lab_host="DH10B"
/clone_lib="Gm-c1045"
/note="Vector: pBluescriptII SK+; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from etiolated hypocotyl tissue of 9-10 day old seedlings of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) primer with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and XhoI. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

ORIGIN

Query Match 43.8%; Score 402; DB 10; Length 630;
Best Local Similarity 86.6%; Pred. No. 1.5e-68;
Matches 471; Conservative 0; Mismatches 60; Indels 13; Gaps 2;

```
Qy          61 TTTGACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGG 120
               ||||| ||| |||||| || ||  |||||| |||  || ||||||||||||||||||||
Db           7 TTTGATCTGATATAGTGAATAGAAATGGCTAAAAACGAGTTGAGGCTTTTGGGTGCTTGG 66

Qy         121 TTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 180
               |||||||||| |||||||||||||||||||||||||||||||||||| | |||||||||
Db          67 TTCAGTCCATATGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTCTAGATTATGAG 126

Qy         181 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240
               |||||||||||||| ||||||||||||||||  ||||||||||||||||||||||||||
Db         127 GTTGTTGAAGAGACCTTGAATCCCAAAAGTGATCTGCTTCTTAAGTCCAACCCTGTGCAC 186

Qy         241 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 300
               |||||||||||||| |  ||||||||||||||||||||||||||||||||||||||||
Db         187 AAGAAAATCCCAGTTCTACTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTA 246

Qy         301 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 360
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
```

```
Db        247 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCTTAT 306

Qy        361 GATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 420
              || ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||| |
Db        307 GACCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGTATACGTCCATA 366

Qy        421 AAAAGTGTTCTAGCGACTG------------AAGATGATGAGGCAAAGAAGCTACACTTT 468
              | | |  | |    | |||            |||||||||||||||||||| |||||||
Db        367 AGGAATACTATTTTGGCTGAAGCGGCTGATCAAGATGATGAGGCAAAGAAGCCACACTTT 426

Qy        469 GAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGG 528
              | |  |  ||||||||  ||||||| | ||||||||||||||||||||||||||||||||
Db        427 GTGAGAATGGAAGAAGCTCTTGAGAGGATGGAAGAAGTGTTCAACAAGTGCAGTGAAGGA 486

Qy        529 AAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTG 588
              | ||| ||||| |||||||||  | ||||||  || |||||||||  |||||||  | ||  |
Db        487 AGGGCTTATTTTGGAGGAGATTCAATTGGA-TTATTGACATTGTTTTTGGAGCCCTTGGG 545

Qy        589 AGTT 592
                |||
Db        546 GGTT 549


RESULT 15
BG653588
LOCUS       BG653588                 502 bp    mRNA    linear   EST 29-NOV-2001
DEFINITION  sad76c09.y1 Gm-c1051 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-c1051-6257 5' similar to TR:Q9XIF8 Q9XIF8 F23H11.1 PROTEIN.
            ;, mRNA sequence.
ACCESSION   BG653588
VERSION     BG653588.1  GI:13790997
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 502)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            High quality sequence stop: 416.
FEATURES             Location/Qualifiers
     source          1. .502
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1051-6257"
                     /tissue_type="floral meristematic mRNA"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1051"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from floral
                     meristematic mRNA provided by Dr. Halina Knap of Clemson
                     University. Complementary DNA was synthesized from mRNA
                     using a primer consisting of a poly(dT) sequence with a
                     XhoI restriction site. EcoRI adapters were ligated to the
                     blunt-ended cDNA fragments followed by XhoI digestion. The
                     cDNA fragments were directionally cloned into the
                     EcoRI-XhoI restriction site of the pBluescript vector. The
                     ligated cDNA fragments were transformed into DH10B host
                     cells (GibcoBRL). This library was constructed in the
                     laboratory of Dr. Randy Shoemaker."
ORIGIN

  Query Match             41.7%;  Score 382.8;  DB 12;  Length 502;
  Best Local Similarity   87.2%;  Pred. No. 9e-65;
  Matches  436;  Conservative    0;  Mismatches   52;  Indels   12;  Gaps    1;

Qy        115 GCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGAT 174
              |||||||||||||||| |||||||||||||| |||||||||||||||||||| | |||
Db          2 GCTTGGTTCAGTCCATATGCCCTGAGGGTGCAAATTGCCCTTAACCTCAAGGGTCTAGAT 61

Qy        175 TATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCT 234
              |||||||||||||||||||| ||||||||||||||||  ||||||||||||||||||||
Db         62 TATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAGTGATCTGCTTCTTAAGTCCAACCCT 121

Qy        235 GTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATC 294
              |||||||||||||||||||| |  |||||||||||||||||||||||||||||||||||
Db        122 GTGCACAAGAAAATCCCAGTTCTACTCCATGGAGATAAAGTCATATGTGAATCTGCAATC 181

Qy        295 ATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAAT 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAAT 241

Qy        355 GCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACG 414
              || ||||| ||||||||||||||||||||||||||||||||||||||||||||  ||||
Db        242 GCTTATGACCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGTATACG 301

Qy        415 TCCTTGAAAAGTGTTCTAGCGACTG------------AAGATGATGAGGCAAAGAAGCTA 462
              ||| | |  | |  | |   | |||            |||||||||||||||||||| |
Db        302 TCCATAAGGAATACTATTTTGGCTGAAGCGGCTGATCAAGATGATGAGGCAAAGAAGCCA 361

Qy        463 CACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGT 522
              ||||||| |  |  |||||||  ||||||| | ||||||||||||||||||||||||||
Db        362 CACTTTGTGGGAATGGAAGAAGCTCTTGAGAGGATGGAAGAAGTGTTCAACAAGTGCAGT 421

Qy        523 GAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGC 582
              ||||| | ||| ||||| |||||||||| |||||| || ||||||||| |||||| |||
Db        422 GAAGGAAGGGCTTATTTTGGAGGAGATACAATTGGAATTATTGACATTGCTTTTGGGAGC 481

Qy        583 TTTTTGAGTTTCATTAGAGT 602
              ||| | |||   | |||||
Db        482 CTTTGGGGTTGGGTGAGAGT 501


Search completed: October 13, 2004, 05:50:02
Job time : 2952 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 13, 2004, 05:00:46 ; Search time 534 Seconds
(without alignments)
8715.098 Million cell updates/sec

Title: US-10-088-945A-14
Perfect score: 918
Sequence: 1 cataaaactccacatttcct.....aatatagtaaaaaaaaaaaa 918

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	887	96.6	1050	13	US-10-424-599-51507	Sequence 51507, A
2	623.2	67.9	1265	13	US-10-424-599-57938	Sequence 57938, A
3	288	31.4	896	9	US-09-765-213A-5	Sequence 5, Appli
4	270	29.4	998	9	US-09-765-213A-1	Sequence 1, Appli
5	268.4	29.2	947	9	US-09-765-213A-3	Sequence 3, Appli
6	249.2	27.1	671	17	US-10-021-323-17539	Sequence 17539, A
7	240	26.1	684	9	US-09-938-842A-2676	Sequence 2676, Ap
8	240	26.1	684	11	US-09-938-842A-2676	Sequence 2676, Ap
9	232.2	25.3	705	9	US-09-938-842A-2658	Sequence 2658, Ap
10	232.2	25.3	705	11	US-09-938-842A-2658	Sequence 2658, Ap
c 11	228.2	24.9	817	9	US-09-770-445-743	Sequence 743, App
c 12	222.8	24.3	779	16	US-10-260-238-4937	Sequence 4937, Ap
13	220	24.0	624	17	US-10-021-323-8555	Sequence 8555, Ap
14	212.4	23.1	606	17	US-10-021-323-499	Sequence 499, App
15	206.6	22.5	589	17	US-10-021-323-15323	Sequence 15323, A
16	202.2	22.0	566	17	US-10-021-323-14067	Sequence 14067, A
17	201.2	21.9	552	17	US-10-021-323-14839	Sequence 14839, A
18	193.6	21.1	513	9	US-09-938-842A-1665	Sequence 1665, Ap
19	193.6	21.1	513	11	US-09-938-842A-1665	Sequence 1665, Ap
c 20	185.2	20.2	578	17	US-10-021-323-17631	Sequence 17631, A
21	181.8	19.8	684	9	US-09-938-842A-695	Sequence 695, App
22	181.8	19.8	684	11	US-09-938-842A-695	Sequence 695, App
c 23	181.4	19.8	824	9	US-09-770-445-717	Sequence 717, App
24	181.2	19.7	971	17	US-10-437-963-50534	Sequence 50534, A
25	177.4	19.3	1128	17	US-10-437-963-52368	Sequence 52368, A
26	175.2	19.1	708	17	US-10-437-963-45535	Sequence 45535, A
27	172	18.7	743	17	US-10-767-701-7710	Sequence 7710, Ap
28	171	18.6	1121	17	US-10-437-963-48754	Sequence 48754, A
29	169.2	18.4	926	13	US-10-425-114-13883	Sequence 13883, A
30	169	18.4	980	13	US-10-425-114-5225	Sequence 5225, Ap
31	167.4	18.2	1067	13	US-10-425-114-28110	Sequence 28110, A
32	165	18.0	809	13	US-10-425-114-19959	Sequence 19959, A
33	163.8	17.8	333	13	US-10-424-599-76241	Sequence 76241, A
34	162.2	17.7	994	17	US-10-437-963-45541	Sequence 45541, A
35	161.2	17.6	921	17	US-10-437-963-29311	Sequence 29311, A
36	161.2	17.6	1004	17	US-10-437-963-84306	Sequence 84306, A
37	159.6	17.4	513	17	US-10-021-323-13431	Sequence 13431, A
38	157.6	17.2	851	13	US-10-425-114-24808	Sequence 24808, A
39	154.4	16.8	903	13	US-10-425-114-22273	Sequence 22273, A
40	154.2	16.8	998	17	US-10-437-963-6569	Sequence 6569, Ap
41	153.8	16.8	910	13	US-10-425-114-26127	Sequence 26127, A
42	153.8	16.8	1028	17	US-10-767-701-13603	Sequence 13603, A
c 43	153.6	16.7	516	17	US-10-021-323-16335	Sequence 16335, A
44	153.6	16.7	708	17	US-10-437-963-4127	Sequence 4127, Ap
45	151.4	16.5	1146	17	US-10-437-963-41800	Sequence 41800, A

ALIGNMENTS

RESULT 1
US-10-424-599-51507
; Sequence 51507, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 51507
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17524C.1
US-10-424-599-51507

Query Match 96.6%; Score 887; DB 13; Length 1050;
Best Local Similarity 98.7%; Pred. No. 1.1e-220;
Matches 905; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

```
Qy          1 CATAAAACTCCACATTTCCTGCTGAGTAACCTAACAAAACAAACACAATATTGCTCCGTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db        128 CATAAAACTCCACATTTCCTGCTGAGTAACCTAACAAAACAAACACAGCATTGCTCCGTG 187

Qy         61 TTTGACCTGTTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGG 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        188 TTTGACCTGTTATAGTAAACAGTGATGGCTCAAAGGGACTTGAGGCTTTTGGGTGCTTGG 247

Qy        121 TTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 180
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        248 TTCAGTCCATTTGTCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAG 307

Qy        181 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 GTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 367

Qy        241 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 AAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTT 427

Qy        301 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATAT 487

Qy        361 GATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 420
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 GATGGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTG 547

Qy        421 AAAAGTGTTCTAGCGACTGAA--GATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGG 478
              |||||||||||  | |||||  ||||||||||||||||||||||||||||||||||||
Db        548 AAAAGTGTTCTACTGGCTGAAGTGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGG 607

Qy        479 AAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATT 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 AAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATT 667

Qy        539 TCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTA 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 TCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTA 727

Qy        599 GAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGA 658
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        728 GAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGCACCCTGGTTTGA 787

Qy        659 CCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTG 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 CCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTG 847

Qy        719 AAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAA 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        848 AAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAAATGGGCTGCTGCAGCTGCTGCAA 907

Qy        779 AGTAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTA 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 AGTAAATGGAATCAAATTAATTGCTGGATGAATTTCAAAAATTGTTGTGCAAGTTATTTA 967

Qy        839 TATCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGAT 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 TATCTGAGGCTATGTTTGTTGCAACTTTATATATTTAAAAGTCAAAATAAATGTTATGAT 1027

Qy        899 AATATAGTAAAAAAAAA 915
              ||||||||||||||| |
Db       1028 AATATAGTAAAAAAATA 1044
```

```
RESULT 2
US-10-424-599-57938
; Sequence 57938, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 57938
;   LENGTH: 1265
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_2332C.1
US-10-424-599-57938

  Query Match            67.9%;  Score 623.2;  DB 13;  Length 1265;
  Best Local Similarity  85.2%;  Pred. No. 8e-152;
  Matches  758;  Conservative    0;  Mismatches  118;  Indels   14;  Gaps    5;

Qy         30 CCTAACAAAACAAACACAATATTGCTCCGTGTTTGACCTGTTATAGTAAACAGTGATGGC 89
              || |||||| |||||||||  |||| || |  |  | ||  |       |     |
Db         46 CCCAACAAAGCAAACACAACTTTGCCCCCTTGTAAAACTTCTTATTGTGATGTCTAAAAG 105

Qy         90 TGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGAT 149
               |||   ||||||| |||||||| |  |||||||| ||||||||||||||||||||||
Db        106 CGAA---GACTTGAAGCTTTTGGGAGGCTGGTTCAGCCCATTTGCCCTGAGGGTGCAGAT 162

Qy        150 TGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAG 209
              |||||||||||||||||| | || ||||||||||||||||||| ||||||||||||||
Db        163 TGCCCTTAACCTCAAGGGTCTAGAATATGAGGTTGTTGAAGAGACCTTGAATCCCAAAAG 222

Qy        210 TGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGA 269
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 TGACCTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGA 282

Qy        270 TAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAA 329
              |||||||| |||||||||||||||||||||||||||||| ||||||| ||||   | ||
Db        283 TAAAGTCATTTGTGAATCTGCAATCATAGTTGAGTACATTGATGAGGCTTGG---ACTAA 339

Qy        330 TGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTC 389
              || || |||||||||||||||||||| |||||||| |||||||| ||||||||  || |
Db        340 TGTTCCCTCCATCCTTCCACAAAATGCTTATGATCGTGCTAATGCTCGATTTTGGTTTGC 399

Qy        390 TTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGA 449
              |||||  ||||| |||||| |||||||||||| |||||||||||  | ||||||||||||
Db        400 CTACATTGATGAGAAGTGGTTTACGTCCTTGAGAAGTGTTCTAGTGGCTGAAGATGATGA 459

Qy        450 GGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTT 509
              |||||||||||  |||||||||||||||  |||||||  ||||||||  | ||||||||||||
Db        460 GGCAAAGAAGCCACACTTTGAGCAAGCAGAAGAAGGGCTTGAGAGGTTGGAAGAAGTGTT 519

Qy        510 CAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACAT 569
              |||||||| ||||||||||||||||||||||||||||||  ||||||||  |||||||
Db        520 CAACAAGTACAGTGAAGGGAAGGCCTATTTCGGAGGAGATAGCATTGGATTCATTGACAT 579

Qy        570 TGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATT 629
              |||||||| ||||| |||||||  || ||||||  |||| | |||| || |||||||||
Db        580 TGGTTTTGGGAGCTTCTTGAGTTGGATGAGAGTCATAGAGGAGATGAGTGGAAGAAAATT 639

Qy        630 GCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCC 689
              |||||||||| |||| |||||||||||||| |||||||||||| ||||||||||||||
Db        640 GCTTGATGAAAAGAAGCACCCTGGTTTGACCCAATGGGCTGAAACGTTTGCTGCTGATCC 699

Qy        690 TGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCA 749
              |||||||||||| |||| |||||||||| |||||||||||||||| ||||||||||
Db        700 TGCTGTGAAGGGCATTCTTCCAGAGACTGATAAGCTTGTTGAGTTTGCCAAGATTCTTCA 759

Qy        750 GCTAAAATGGGCTGCTGCA------GCTGCTGCAAAGTAAATGGAATCAAATTAATTGCT 803
              |||||||||| ||||||||      || |||||||||||||||||||||||||||||||
Db        760 GCTAAAATGGACTGCTGCAGCAGCTGCAGCTGCAAAGTAAATGGAATCAAATTAATTGCG 819

Qy        804 GGATGAATTTCAAAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAAC 863
              || | ||||  |||||||||||| |||||| ||| ||||| |||||||| ||||||||||
Db        820 AGA-GTATTTTCAAAATTGTTGTCCAAGTTGTTTTTATCTCAGGCTATG-TTGTTGCAAC 877

Qy        864 TTTATATATTTAAAAGTCAAAATAAATGTTATGATAATATAGTAAAAAAA 913
              ||||| |||||||||||  |   ||||| | |     | |   | ||| |||
Db        878 TTTATTTATTTAAAAGTTATTTTAAATTTAAAATGTAAAATATTAAGAAA 927
```

```
RESULT 3
US-09-765-213A-5
; Sequence 5, Application US/09765213A
```

; Patent No. US20020079846A1
; GENERAL INFORMATION:
; APPLICANT: Facchini, Peter J
; TITLE OF INVENTION: No. US20020079846A1el Glutathione-S-Transferase Nucleic Acids and
; TITLE OF INVENTION: Polypeptides and Methods of Use
; FILE REFERENCE: 22542-001
; CURRENT APPLICATION NUMBER: US/09/765,213A
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/176708
; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Papaver somniferum
US-09-765-213A-5

Query Match 31.4%; Score 288; DB 9; Length 896;
Best Local Similarity 59.6%; Pred. No. 1.8e-64;
Matches 486; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

```
Qy         96 GGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCT 155
Db         81 GGAGGTAAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACT 140

Qy        156 TAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATT 215
Db        141 CAACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACT 200

Qy        216 GCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGT 275
Db        201 TCTTCTGAAATCAAATCCTATTTACAAGAAAATTCCTGTTATGATTCATGGTGATAAACC 260

Qy        276 CATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCT 335
Db        261 CATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTTTGGGCTTCTGCTGGACA 320

Qy        336 CTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACAT 395
Db        321 TTCTATCATCCCGTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACAT 380

Qy        396 CGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAA 455
Db        381 TGATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAA 440

Qy        456 GAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAA 515
Db        441 AAAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTGGAAGAAGCTTATCAGAA 500

Qy        516 GTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTT 575
Db        501 AACTAGTAAAGGAAAAGACTTTTTCGGGGAAGAAAAAATTGGATACATTGATATTGCATT 560

Qy        576 TGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGA 635
Db        561 TGGGTGTTATATAGGTTGGATTAGAGTTACAGAGAAAATGAATGGAATCAAACTATTTGA 620

Qy        636 TGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGT 695
Db        621 TGAAACAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACAGT 680

Qy        696 GAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTCAGCTAAA 755
Db        681 TAAATCTGTTATGCCTGAAACTGATGCTCTCATGGAGTTTGCTAAGAAGATCTTTGGATC 740

Qy        756 ATGGGCTGCTGCAGCTGCTGCAAAGTAAATGGAATCAAATTAATTGCTGGATGAATTTCA 815
Db        741 TAAGCCTCCTCCTTCAAACTAGAAAAAGTTGTTAACAATGAAATATCTTAGAGATGTTTA 800

Qy        816 AAAATTGTTGTGCAAGTTATTTATATCTGAGGCTATGTTTGTTGCAACTTTATATATTTA 875
Db        801 AGCTTTGTGTTTGTTTTTTCAGTGTTGTGTAGCAATGCTTAAGAACTGTTTGTAGAAATG 860

Qy        876 AAAGTCAAAATAAATGTTATGATAATATAGTAAAAA 911
Db        861 ATCAAGAACAGTAGCTGTAAAAAAAAAAAAAAAAA 896
```

RESULT 4
US-09-765-213A-1
; Sequence 1, Application US/09765213A
; Patent No. US20020079846A1
; GENERAL INFORMATION:
; APPLICANT: Facchini, Peter J
; TITLE OF INVENTION: No. US20020079846A1el Glutathione-S-Transferase Nucleic Acids and
; TITLE OF INVENTION: Polypeptides and Methods of Use
; FILE REFERENCE: 22542-001
; CURRENT APPLICATION NUMBER: US/09/765,213A
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/176708
; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Papaver somniferum
US-09-765-213A-1

Query Match 29.4%; Score 270; DB 9; Length 998;
Best Local Similarity 63.6%; Pred. No. 9.3e-60;
Matches 411; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```
Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
Db         99 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 158

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
Db        159 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 218

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
Db        219 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 278

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
Db        279 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 338

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
Db        339 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 398

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
Db        399 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 458

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
Db        459 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 518

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
Db        519 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 578

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
Db        579 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 638

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
Db        639 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCTGATGAGACAGTT 698

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742
```

```
Db        699 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 744


RESULT 5
US-09-765-213A-3
; Sequence 3, Application US/09765213A
; Patent No. US20020079846A1
; GENERAL INFORMATION:
;  APPLICANT: Facchini, Peter J
;  TITLE OF INVENTION: No. US20020079846A1el Glutathione-S-Transferase Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides and Methods of Use
;  FILE REFERENCE: 22542-001
;  CURRENT APPLICATION NUMBER: US/09/765,213A
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 60/176708
;  PRIOR FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 947
;   TYPE: DNA
;   ORGANISM: Papaver somniferum
US-09-765-213A-3

  Query Match            29.2%;  Score 268.4;  DB 9;  Length 947;
  Best Local Similarity  63.5%;  Pred. No. 2.4e-59;
  Matches  410;  Conservative    0;  Mismatches  236;  Indels    0;  Gaps    0;

Qy         97 GACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTT 156
Db         48 GAGGTGAAGATTTTAGGTGGATGGCCAAGTCCATTTGTGATGAGGCCTAGAATTGCACTC 107

Qy        157 AACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAAAGTGAATTG 216
Db        108 AACATTAAATCAGTCAAGTATTATCTTCTTGAAGAGACATTTGGTAGCAAAAGTGAACTT 167

Qy        217 CTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAGATAAAGTC 276
Db        168 CTTCTGAAATCAAATCCTATTTACAAGAAGATGCCTGTCTTGATTCACGGTGATAAACCC 227

Qy        277 ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTC 336
Db        228 ATCTGTGAATCAATGATCATTGTTCAGTACATTGATGATGTCTGGGCTTCTGCTGGTCAT 287

Qy        337 TCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATC 396
Db        288 TCCATCATCCCTTCTGATCCTTATGATGCTTCCATTGCTCGTTTCTGGGCAACCTACATT 347

Qy        397 GATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAG 456
Db        348 GATGACAAGTTCTTTCCGTCTTTAATGGGGATTGCAAAGAGTAAGGATGCAGAAGAAAAA 407

Qy        457 AAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAG 516
Db        408 AAAGCAGCCATTGAACAGGCGATTGCAGCTTTTGGTATACTTGAAGAAGCTTATCAGAAA 467

Qy        517 TGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTT 576
Db        468 ACTAGTAAAGGAAAAGATTTTTTCGGTGGAGAAAAAATTGGGTATGTCGATATTGCATTT 527

Qy        577 GGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGAT 636
Db        528 GGGTGTTATGTTGGCTGGATTAGAGTTACAGAGAAGATGAACGGAATCAAACTATTTGAT 587

Qy        637 GAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTG 696
Db        588 GAAGAAAAAGTTCCAGGGCTTACAAAATGGGCTGAGAAATTTTGTGCAGATGAGACGGTT 647

Qy        697 AAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGA 742
Db        648 AAATCTGTTATGCCTGAAACTGATGCCCTCATGGAGTTTGCTAAGA 693


RESULT 6
US-10-021-323-17539
; Sequence 17539, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
;  APPLICANT: Deikman, Jill
;  APPLICANT:  Feng, Paul C.C.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Ziegler, Todd E.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(52274)B
;  CURRENT APPLICATION NUMBER: US/10/021,323
;  CURRENT FILING DATE:  2001-12-12
;  PRIOR APPLICATION NUMBER: US 60/255, 619
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 17539
;   LENGTH: 671
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(671)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: LIB3831-001-Q1-K6-B3
US-10-021-323-17539

  Query Match            27.1%;  Score 249.2;  DB 17;  Length 671;
  Best Local Similarity  65.2%;  Pred. No. 2e-54;
  Matches  431;  Conservative    0;  Mismatches  219;  Indels   11;  Gaps    4;

Qy        124 AGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTT 183
Db          5 AGTCCATATGCGATGAGGCCAAGNATTGCCCTTAACCTCAAATCCCTCACTTACGAATAC 64

Qy        184 GTTGAAGAGACTTT---GAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCAC 240
Db         65 ATTGAAGAGAAGTTATGGGAAGGCAAAAGCGAGCTTCTTCTCCAATCAAACCCTGTTTAC 124

Qy        241 AAGAAAATCCCAGTTTTCTTCCATGGAGAT---AAAGTCATATGTGAATCTGCAATCATA 297
Db        125 AAGAAAATCCCAGTCCTCATTCACGGCGATAACAAGCCAATCTGCGAATCTCTCATCATC 184

Qy        298 GTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCA 357
Db        185 GTACAATACATCGACGAGGCTTGGTCTTCC---GGTCCTTCTATTCTTCCTTCTGATCCC 241

Qy        358 TATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCC 417
Db        242 TGTGAACGTGCTGTCGCTCGGTTTTGGGCTGCATATCTTGACGACAAGTGGTTCCCAGCG 301

Qy        418 TTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCG 477
Db        302 GTGAAAAGTATAGGAGTCGCTAAAGGAGAGGAAGCAAAGAAAGCAGCAATAGCGCAAGTG 361

Qy        478 GAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTAT 537
Db        362 GAAGCAGGGTTGGCTTTGATGGAAGAAGCATACGACAAGTGCAGCAAAGGGAAGGGTTAC 421

Qy        538 TTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATT 597
Db        422 TTTGGTGGGGATGAAGTTGGGTACCTTGATATAGCTTTTGGGAGCTTCTTGGGATGGCTT 481

Qy        598 AGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTG 657
Db        482 AGAGTGACGGAGAAGTTCAATGCGATGAAGCTGCTGGACGAGGGGAAGACGCCTGGTCTG 541

Qy        658 ACCCTATGGGCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACT 717
Db        542 GTAAAATGGGCGGACAGGTTCTGTTCCCATGCCGCTGTGAAGGATGTTATGCCTGATACT 601

Qy        718 GAAAAGCTTGTTGAGTTTGCAAAGATTCT--TCAGCTAAAATGGGCTGCTGCAGCTGCTG 775
```

```
Db        602 GACAAGCTTGCGGAGTTTGGTAAGTTTGTTATTGCCAAAATAGGGCCGCTGGCGCTGCTG 661

Qy        776 C 776
              |
Db        662 C 662
```

RESULT 7
US-09-938-842A-2676
; Sequence 2676, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2676
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2676

Query Match 26.1%; Score 240; DB 9; Length 684;
Best Local Similarity 59.6%; Pred. No. 5e-52;
Matches 405; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

```
Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
              |||||      |||| ||| |||  | ||  | |||   |||    |  | | ||||||  |
Db          1 ATGGCGACCGAGGACGTGAAGCTGATCGGCTCATGGGCGAGTGTTTACGTCATGAGGGCG 60

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
                ||| || ||  |||||||    | |     || ||  |  || | |||||  |    | |
Db         61 AGGATCGCTCTCCACCTCAAATCTATTAGCTACGAATTCCTTCAGGAGACGTATGGTTCA 120

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
              || || |||||||| || || || ||||| || |||||||| || || ||| |   | ||
Db        121 AAGAGCGAATTGCTCCTCAAATCAAACCCGGTTCACAAGAAGATGCCGGTTCTGATTCAC 180

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
              |  || |||    | || || ||    ||||| ||| | || || || |||| ||||  |
Db        181 GCTGACAAACCGGTGTGCGAGTCCAACATCATCGTTCATTATATCGACGAGGCTTGGAAC 240

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
                 ||  |  ||||| || ||     || |||| || || || | ||| || ||||||
Db        241 TCTTCTGGACCTTCCATTCTCCCGTCCCATCCATACGACCGGGCCATTGCTCGGTTTTGG 300

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
              | | | ||||| || ||  ||||| |||   ||  ||| |||| | ||| |   || |||
Db        301 GCTGCCTACATAGACGATCAGTGGTTTATCTCTGTGAGAAGTATCCTAACAGCTCAAGGA 360

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
              || || |  |||||   |  | | |  ||||  ||||||  |    ||    ||||  ||
Db        361 GACGAAGAGAAGAAAGCAGCCATAGCTCAAGTTGAAGAAAGGACCAAGCTTCTGGAGAAA 420

Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
              |  |||||| | || ||  |||| ||  | |  |||   || ||    || || |   |
Db        421 GCATTCAACGATTGTAGCCAAGGAAAACCGTTCTTCAACGGTGACCATATCGGTTACCTC 480

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
              |||||||  || || ||||| ||  |||     ||||||   |||          |
```

```
Db        481 GACATTGCCTTGGGGAGCTTCTTAGGTTGGTGGAGAGTCGTCGAGTTGGATGCCAATCAT 540

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
              ||||| |||||||| || ||    ||   | |   |  ||||| ||    ||   || | |
Db        541 AAATTTCTTGATGAGACCAAAACTCCCTCTCTAGTCAAATGGGCAGAGCGGTTCTGTGAT 600

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATT 744
              |||||  |||||||      |  |||| |||| |  |||||| | ||| || |||| ||
Db        601 GATCCCGCTGTGAAACCTATAATGCCCGAGATTACAAAGCTCGCTGAATTCGCAAGGAAG 660

Qy        745 CTTCAGCTAAAATGGGCTGC 764
              ||   |  ||  ||  ||
Db        661 CTCTTTCCTAAGCGGCAAGC 680
```

RESULT 8
US-09-938-842A-2676
; Sequence 2676, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2676
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2676

Query Match 26.1%; Score 240; DB 11; Length 684;
Best Local Similarity 59.6%; Pred. No. 5e-52;
Matches 405; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

```
Qy         85 ATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTG 144
              |||||      |||| ||| |||  | ||  | |||   |||    |  | | ||||||  |
Db          1 ATGGCGACCGAGGACGTGAAGCTGATCGGCTCATGGGCGAGTGTTTACGTCATGAGGGCG 60

Qy        145 CAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCC 204
                ||| || ||  |||||||    | |     || ||  |  || | |||||  |    | |
Db         61 AGGATCGCTCTCCACCTCAAATCTATTAGCTACGAATTCCTTCAGGAGACGTATGGTTCA 120

Qy        205 AAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCAT 264
              || || |||||||| || || || ||||| || |||||||| || || ||| |   | ||
Db        121 AAGAGCGAATTGCTCCTCAAATCAAACCCGGTTCACAAGAAGATGCCGGTTCTGATTCAC 180

Qy        265 GGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCC 324
              |  || |||    | || || ||    ||||| ||| | || || || |||| ||||  |
Db        181 GCTGACAAACCGGTGTGCGAGTCCAACATCATCGTTCATTATATCGACGAGGCTTGGAAC 240

Qy        325 AACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGG 384
                 ||  |  ||||| || ||     || |||| || || || | ||| || ||||||
Db        241 TCTTCTGGACCTTCCATTCTCCCGTCCCATCCATACGACCGGGCCATTGCTCGGTTTTGG 300

Qy        385 GTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGAT 444
              | | | ||||| || ||  ||||| |||   ||  ||| |||| | ||| |   || |||
Db        301 GCTGCCTACATAGACGATCAGTGGTTTATCTCTGTGAGAAGTATCCTAACAGCTCAAGGA 360

Qy        445 GATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAA 504
              || || |  |||||   |  | | |  ||||  ||||||  |    ||    ||||  ||
Db        361 GACGAAGAGAAGAAAGCAGCCATAGCTCAAGTTGAAGAAAGGACCAAGCTTCTGGAGAAA 420
```

```
Qy        505 GTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTT 564
              |  |||||| | || ||  |||| ||  | |  |||   || ||    || || |   |
Db        421 GCATTCAACGATTGTAGCCAAGGAAAACCGTTCTTCAACGGTGACCATATCGGTTACCTC 480

Qy        565 GACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGA 624
              |||||||  || || |||||  ||  |||     ||||||   |||           |
Db        481 GACATTGCCTTGGGGAGCTTCTTAGGTTGGTGGAGAGTCGTCGAGTTGGATGCCAATCAT 540

Qy        625 AAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCT 684
              |||||  |||||||| || ||    ||   | |   |  |||||| ||    ||   || |
Db        541 AAATTTCTTGATGAGACCAAAACTCCCTCTCTAGTCAAATGGGCAGAGCGGTTCTGTGAT 600

Qy        685 GATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATT 744
              |||||  ||||||||      |  |||| |||| |  ||||||  | ||| || |||| ||
Db        601 GATCCCGCTGTGAAACCTATAATGCCCGAGATTACAAAGCTCGCTGAATTCGCAAGGAAG 660

Qy        745 CTTCAGCTAAAATGGGCTGC 764
              ||    |  ||  ||   ||
Db        661 CTCTTTCCTAAGCGGCAAGC 680


RESULT 9
US-09-938-842A-2658
; Sequence 2658, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2658
;   LENGTH: 705
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2658

  Query Match            25.3%;  Score 232.2;  DB 9;  Length 705;
  Best Local Similarity   59.7%;  Pred. No. 5.5e-50;
  Matches  409;  Conservative    0;  Mismatches  273;  Indels    3;  Gaps    1;

Qy         92 AAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTG 151
              ||  |||  |||  || || || |  |||| ||||||||  ||| | ||     |||| |
Db         11 AAGAGGAAGTGAAACTATTAGGAGTATGGTACAGTCCATACGCCATAAGACCTAAGATCG 70

Qy        152 CCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATC---CCAAAA 208
              | |||   |||||     ||||||||||     ||||| || | | ||  |    |||| |
Db         71 CTCTTCGTCTCAAATCAGTGGATTATGATTACGTTGAGGAAAATCTGTTTGGATCCAAGA 130

Qy        209 GTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAG 268
              ||||  | ||||| || || ||||| || |||||||||| |||| ||| || ||||
Db        131 GTGAGCTTCTTCTCAAATCAAACCCGGTTCACAAGAAAGTCCCTGTTCTCCTCCACAACA 190

Qy        269 ATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACA 328
              | |||   ||    || |||   | ||| ||||| ||||| ||||||   |||  |
Db        191 ACAAACCGATTGTCGAGTCTCTCAACATCGTTGAATACATTGATGAGACGTGGAACTCAT 250

Qy        329 ATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTT 388
              ||| |  ||||| |||||   | || | |||||||| |||  ||| || ||||||  |
Db        251 CTGCACCGTCCATTCTTCCTTCACATCCTTATGATCGTGCTCTTGCTCGCTTTTGGTCTG 310

Qy        389 CTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATG 448
              | | | ||| |||||||| || || | |||| ||  |         ||| ||   || |
Db        311 ACTTCGTTGATAACAAGTGGTTTCCGGCGTTGAGAATGGCGGCGATCACTAAATCGGAAG 370

Qy        449 AGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGT 508
              | |||||   |  |  | | ||  ||| |||||||| | |   | |  | || || | ||
Db        371 ACGCAAAAGCGAAAGCCATGGAAGAAGTGGAAGAAGGGTTGTTGCAACTCGAGGATGCGT 430

Qy        509 TCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACA 568
              |           ||  |||||||  ||| |||||| || ||   | || || ||| | || |
Db        431 TTGTTTCTATAAGCAAAGGGAAACCCTTTTTCGGCGGTGAAGCAATCGGGTTTATGGATA 490

Qy        569 TTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAAT 628
              || | |||||||||||| ||  | || | | ||    ||| ||  | || | |  |||
Db        491 TTTGCTTTGGAAGCTTTGTGGTTCTCTTGAAAGCTAGAGAAAAGTTTAAAGCAGAAAAGC 550

Qy        629 TGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATC 688
              |  | || ||| | ||     |||  | |       ||||||  ||     ||      | |||
Db        551 TTTTAGACGAATCAAAAACTCCTTCTCTTTGTAAATGGGCCGACCGGTTCTTGTCCGATG 610

Qy        689 CTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTC 748
              | ||||||     |    || ||||  |||||  | | |||||||  | |     |||
Db        611 AAACGGTGAAGAATGTGGCACCGGAGATCGAAAAAGTAGCTGAGTTTCTACAAGAACTTG 670

Qy        749 AGCTAAAATGGGCTGCTGCAGCTGC 773
              || | | |       | |||||| |
Db        671 AGGTTAGAGCTCAATCCGCAGCTTC 695


RESULT 10
US-09-938-842A-2658
; Sequence 2658, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2658
;   LENGTH: 705
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2658

  Query Match            25.3%;  Score 232.2;  DB 11;  Length 705;
  Best Local Similarity   59.7%;  Pred. No. 5.5e-50;
  Matches  409;  Conservative    0;  Mismatches  273;  Indels    3;  Gaps    1;

Qy         92 AAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTG 151
              ||  |||  |||  || || || |  |||| ||||||||  ||| | ||     |||| |
Db         11 AAGAGGAAGTGAAACTATTAGGAGTATGGTACAGTCCATACGCCATAAGACCTAAGATCG 70

Qy        152 CCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATC---CCAAAA 208
              | |||   |||||     ||||||||||     ||||| || | | ||  |    |||| |
Db         71 CTCTTCGTCTCAAATCAGTGGATTATGATTACGTTGAGGAAAATCTGTTTGGATCCAAGA 130

Qy        209 GTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGGAG 268
              ||||  | ||||| || || ||||| || |||||||||| |||| ||| || ||||
Db        131 GTGAGCTTCTTCTCAAATCAAACCCGGTTCACAAGAAAGTCCCTGTTCTCCTCCACAACA 190

Qy        269 ATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACA 328
```

```
Db        191 ACAAACCGATTGTCGAGTCTCTCAACATCGTTGAATACATTGATGAGACGTGGAACTCAT 250

Qy        329 ATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTT 388

Db        251 CTGCACCGTCCATTCTTCCTTCACATCCTTATGATCGTGCTCTTGCTCGCTTTTGGTCTG 310

Qy        389 CTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATG 448

Db        311 ACTTCGTTGATAACAAGTGGTTTCCGGCGTTGAGAATGGCGGCGATCACTAAATCGGAAG 370

Qy        449 AGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGT 508

Db        371 ACGCAAAAGCGAAAGCCATGGAAGAAGTGGAAGAAGGGTTGTTGCAACTCGAGGATGCGT 430

Qy        509 TCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACA 568

Db        431 TTGTTTCTATAAGCAAAGGGAAACCCTTTTTCGGCGGTGAAGCAATCGGGTTTATGGATA 490

Qy        569 TTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAAT 628

Db        491 TTTGCTTTGGAAGCTTTGTGGTTCTCTTGAAAGCTAGAGAAAAGTTTAAAGCAGAAAAGC 550

Qy        629 TGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGATC 688

Db        551 TTTTAGACGAATCAAAAACTCCTTCTCTTTGTAAATGGGCCGACCGGTTCTTGTCCGATG 610

Qy        689 CTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCTTC 748

Db        611 AAACGGTGAAGAATGTGGCACCGGAGATCGAAAAAGTAGCTGAGTTTCTACAAGAACTTG 670

Qy        749 AGCTAAAATGGGCTGCTGCAGCTGC 773

Db        671 AGGTTAGAGCTCAATCCGCAGCTTC 695


RESULT 11
US-09-770-445-743/c
; Sequence 743, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
;  APPLICANT:  Gorlach, Jorn
;  APPLICANT:  An, Yong-Qiang
;  APPLICANT:  Hamilton, Carol M.
;  APPLICANT:  Price, Jennifer L.
;  APPLICANT:  Raines, Tracy M.
;  APPLICANT:  Yu, Yang
;  APPLICANT:  Rameaka, Joshua G.
;  APPLICANT:  Page, Amy
;  APPLICANT:  Matthew, Abraham V.
;  APPLICANT:  Ledford, Brooke L.
;  APPLICANT:  Woessner, Jeffrey P.
;  APPLICANT:  Haas, William David
;  APPLICANT:  Garcia, Carlos A.
;  APPLICANT:  Kricker, Maja
;  APPLICANT:  Slader, Ted
;  APPLICANT:  Davis, Keith R.
;  APPLICANT:  Allen, Keith
;  APPLICANT:  Hoffman, Neil
;  APPLICANT:  Hurban, Patrick
;  TITLE OF INVENTION:  Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE:  2023US (PARA-012PRV)
;  CURRENT APPLICATION NUMBER: US/09/770,445
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER:  US 60/178,472
;  PRIOR FILING DATE:  2000-01-27
;  NUMBER OF SEQ ID NOS:  999
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 743
;   LENGTH: 817
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(817)
;   OTHER INFORMATION: n = A,T,C or G
US-09-770-445-743

  Query Match             24.9%;  Score 228.2;  DB 9;  Length 817;
  Best Local Similarity   59.3%;  Pred. No. 6.6e-49;
  Matches  396;  Conservative    0;  Mismatches  269;  Indels    3;  Gaps    1;

Qy         70 TTATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCA 129

Db        817 TTAGACAAAACAAAGATGGTAGAGAGAGAGGTGAAACTGTTGGGAACATGGTACAGTCCG 758

Qy        130 TTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAA 189

Db        757 GTCGTGATAAGAGTAAAGATCGCTCTTCGTCTCAAATCGGTTGATTATGATTACGTTGAA 698

Qy        190 GAGACTTTG---AATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAA 246

Db        697 GAGAACTTGTTCGGATCTAAGAGTGAGCTTCTTCTCAAATCGAACCCGGTTTACAAGAAA 638

Qy        247 ATCCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTAC 306

Db        637 GTGCCGGTTCTCATCCACAGTAATAAATCGGTTTGTGAGTCTCTAAACATCGTTGAATAC 578

Qy        307 ATAGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGA 366

Db        577 ATAGACGAGACGTGGAACTCATCTGGATCGTTCATTCTTCCTTCTCATCCTCATGATCGT 518

Qy        367 GCTAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGT 426

Db        517 GCCCTTGCTCGCTTCTGGTCTGCCTTCGTTGATGACAAGTGGTTTCCGGCTTTGATGGCA 458

Qy        427 GTTCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTG 486

Db        457 GCTGTGGTCGCTAAATCGGAAGACGCAAAAGCAAAAGGCATGGAAGAAGTGGAAGAAGGG 398

Qy        487 CTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGA 546

Db        397 TTGTTGCAACTCGAAGCTGCGTTTATTGCTCTAAGCAAAGGAAAATCTTTTTTCGGCGGT 338

Qy        547 GATACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCA 606

Db        337 GAAACAATCGGGTTCATAGACATTTGCCTTGGAAGCTTTTTGGTTTTCTTGAAAGCTAGA 278

Qy        607 GAGAATATGAATGAAAGAAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGG 666

Db        277 GAGAAGCTTAAAAAAGNNNNNATTNTAGACGAATNNAAAACTCCTTCTCTTTATAGATGG 218

Qy        667 GCTGAAACTTTTGCTGCTGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTT 726

Db        217 GCCAACCGNNNTTTGTCAGATGAAACAGTGAAGAATGTGGTGCCGGAGATAGATAAAGTT 158

Qy        727 GTTGAGTT 734

Db        157 GCCAAGCT 150


RESULT 12
US-10-260-238-4937/c
; Sequence 4937, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, Paul R.
;  APPLICANT:  Moughamer, Todd G.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Cooper, Bret
;  APPLICANT:  Glazebrook, Jane
;  APPLICANT:  Goff, Stephen A.
;  APPLICANT:  Katagiri, Fumiyaki
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Provart, Nicholas
```

; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 4937
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Musa acuminata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (54)..(54)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (58)..(58)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (295)..(295)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (301)..(301)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (398)..(398)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (454)..(454)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (477)..(477)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (484)..(484)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (502)..(502)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (545)..(545)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (606)..(606)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (615)..(615)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (627)..(627)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (642)..(642)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (644)..(644)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (670)..(670)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (711)..(711)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (718)..(718)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (738)..(738)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (743)..(743)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (772)..(772)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-4937

Query Match 24.3%; Score 222.8; DB 16; Length 779;
Best Local Similarity 56.9%; Pred. No. 1.6e-47;
Matches 389; Conservative 0; Mismatches 295; Indels 0; Gaps 0;

```
Qy        87 GGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCA 146
             ||| | | | ||  ||| |||  | || || |||   || || || |  ||| |   |
Db       735 GGCGGCAGGAGAAGTGANGCTGATNGGGGCGTGGCCGAGCCCGTTCGTGCTGCGCCCGAG 676

Qy       147 GATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAA 206
             | | | ||| ||||| |||||  |||| ||  |  |  |  | ||||  ||       ||
Db       675 GGTCGNCCTCAACCTGAAGGGGGTGGAGTACNANTTCCTGCAGGAGAANTTCGGGGAGAA 616

Qy       207 AAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAATCCCAGTTTTCTTCCATGG 266
             || ||  | ||||| |  ||||||||||| ||||||| ||||| ||  || ||||
Db       615 NAGCGAGCTNCTTCTCAGATCCAACCCTGTGTACAAGAAGATCCCCGTCCTCCTCCACCA 556

Qy       267 AGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAA 326
             || ||   | | || || ||    ||||| || |||||| | || | ||| | | | ||
Db       555 CGACAAGCCCNTCTGCGAGTCGATGATCATCGTGGAGTACGTCGACGGGGTCTNGGCTAA 496

Qy       327 CAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGT 386
               |   |||| | ||     | | || || || |       ||| || |||
Db       495 TTCCGGCCAGGNCATCTTNCCTGCCGACCCCTACGAGCGCGNCCTCCACCGCTTCTGGTC 436

Qy       387 TTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGA 446
             |||||||| |||||||| | |||| |    |  |  | |||||   ||
Db       435 CGTCTACATCGACGACAAGTGGTTCCCGTCAATATTCNGCATCGGAAAGGCTGAAACAGA 376

Qy       447 TGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGT 506
             ||||| |||     |  |   |||||||||   |  | |||  ||  | | || || |
Db       375 GGAGGCCAAGGCCGAATCCGCAGAGCAAGCGTGGGCCGGGCTGAAGCTGCTCGAGGAGGC 316
```

```
Qy        507 GTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGA 566
Db        315 GTTTGAGAAGCTGANCAAGGNCAAAGCCTTCTTCGGCGGGGACACCATCGGTTACGTCGA 256

Qy        567 CATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAA 626
Db        255 CATCGCTCTCGGTGCCTACCTGGGATGGGCGAATGTGATCGAGCAGATGACCGGCCTCAA 196

Qy        627 ATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGCTGA 686
Db        195 GTTGTTCGACGAGGAGAAGACACCTCTGCTCGCGGTTTGGGCGGAGAGGTTCTGCGCCCA 136

Qy        687 TCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGATTCT 746
Db        135 CGAGGCCGTGAAGGAGGTGATGCCGGAGACGGAGAAGCTGCTGGAGATCGCCAAGATGCG  76

Qy        747 TCAGCTAAAATGGGCTGCTGCAGC 770
Db         75 TCAGTAGAAATAGGAGTNTCCNGC  52
```

RESULT 13
US-10-021-323-8555
; Sequence 8555, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 8555
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-011-Q1-K6-H9
US-10-021-323-8555

Query Match 24.0%; Score 220; DB 17; Length 624;
Best Local Similarity 64.7%; Pred. No. 7.8e-47;
Matches 391; Conservative 0; Mismatches 205; Indels 8; Gaps 4;

```
Qy        148 ATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGAGACTTTGAATCCCAAA 207
Db          1 ATTGCCCTTAACCTCAATCCCTCACCTTGATACATTGAAG-AGAATTATGGGAAGGCAAA  59

Qy        208 AGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAA-GAAAATCCCAGTTTTCTTCCATGG 266
Db         60 AGCGAGCTTCTTCTCCAATCAAACCCTGTTTACAATGAAAATCCCAGTCCTCATTCACGG 119

Qy        267 AGATAAAGTC---ATATGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTC 323
Db        120 CGATAACAAGCCAATCTGCGAATCTCTCATCATCGTACAATACATCGACGAGGCTTGGTC 179

Qy        324 CAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTG 383
Db        180 TTCC---GGTCCTTCTATTCTTCCTTCTGACCCCTGTGAACGTGCTGTCGCTCGGTTTTG 236

Qy        384 GGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGA 443
Db        237 GGCTGCATATCTTGACGACAAGTGGTTCCCAGCGGTGAAAAGTATAGGAGTCGCTAAAGG 296

Qy        444 TGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGA 503
Db        297 AGAGGAAGCAAAGAAAGCAGCAATAGCGCAAGTGGAAGAAGGGTTGGCTTTGATGGAAGA 356

Qy        504 AGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGT 563
Db        357 AGCATACGACAAGTGCAGCAAAGGGAAGGGTTACTTTGGTGGGGATGAAGTTGGGTACCT 416

Qy        564 TGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAG 623
Db        417 TGATATAGCTTTTGGGAGCTTCTTGGGATGGCTTAGAGTGACGGAGAAGTTCAATGCGAT 476

Qy        624 AAAATTGCTTGATGAAACGAAGTACCCTGGTTTGACCCTATGGGCTGAAACTTTTGCTGC 683
Db        477 GAAGCTGCTGGACGAGGGGAAGACGCCTGGTCTGGTAAAATGGGCGGACAGGTTCTGTTC 536

Qy        684 TGATCCTGCTGTGAAGGGCCTTCTGCCAGAGACTGAAAAGCTTGTTGAGTTTGCAAAGAT 743
Db        537 CCATGCCGCTGTGAAGGACGTTATGCCTGATACCGACAAGCTTGCGGAGTTTGGTAAGTT 596

Qy        744 TCTT 747
Db        597 TGTT 600
```

RESULT 14
US-10-021-323-499
; Sequence 499, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 499
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3825-008-Q1-K6-A7
US-10-021-323-499

Query Match 23.1%; Score 212.4; DB 17; Length 606;
Best Local Similarity 63.7%; Pred. No. 7.3e-45;
Matches 340; Conservative 0; Mismatches 191; Indels 3; Gaps 1;

```
Qy         72 ATAGTAAACAGTGATGGCTGAAAGGGACTTGAGGCTTTTGGGTGCTTGGTTCAGTCCATT 131
Db         67 AGAAGAAAGAAGAATGGCAGAGAGTAACGTTTTGGTATTGGGCACCTGGGCAAGCCCGTT 126

Qy        132 TGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGGGTTTGGATTATGAGGTTGTTGAAGA 191
Db        127 CGTGCTGAGGGTGAAGATTGCACTTCACCTCAAATCTGTCAACTACGAGAACCATGAGGA 186

Qy        192 GACTTT---GAATCCCAAAAGTGAATTGCTTCTTAAGTCCAACCCTGTGCACAAGAAAAT 248
Db        187 AAATTTGCCGGAATCCAAAAGCGACCTTCTTCTCAAATCTAACCCTGTTTATAAGAAAGT 246

Qy        249 CCCAGTTTTCTTCCATGGAGATAAAGTCATATGTGAATCTGCAATCATAGTTGAGTACAT 308
Db        247 CCCAGTCCTCATCCATGGTCATAATCCCATCTGTGAGTCTCTAATCATCGTTGAGTACAT 306

Qy        309 AGATGAGGTTTGGTCCAACAATGCTCTCTCCATCCTTCCACAAAATGCATATGATCGAGC 368
Db        307 CGACGAGGTTTGGACCACCTCTGCTTCTTCCATCCTCCCTTCTGATGCCTATGAACGCGC 366

Qy        369 TAATGCCCGATTTTGGGTTTCTTACATCGATGACAAGTGGCTTACGTCCTTGAAAAGTGT 428
```

```
Db        367 CCAGTCCCGATTTTGGGCGGCCTATGTCGAAGATAAGTTTTCCACTGCCTTAAGAAGAGT 426

Qy        429 TCTAGCGACTGAAGATGATGAGGCAAAGAAGCTACACTTTGAGCAAGCGGAAGAAGTGCT 488

Db        427 GTTGTTTGGTGCTACAGAGGAAGATAAAAGGGCAGCAATGGCCGAAGTCTCAGAAGGGAT 486

Qy        489 TGAGAAGGTGGAAGAAGTGTTCAACAAGTGCAGTGAAGGGAAGGCCTATTTCGGAGGAGA 548

Db        487 GGTGCTGTTGGAGGAAGTCTTTGTGAAATTGAGCAAAGGGAAAGCTTTCTTTGGTGGGGA 546

Qy        549 TACGATTGGATTTGTTGACATTGGTTTTGGAAGCTTTTTGAGTTTCATTAGAGT 602

Db        547 AAATATTGGGTTCGTGGACATAGTGTTCGGTAGCTTGTTGGCGTGGATTGAAGT 600


RESULT 15
US-10-021-323-15323
; Sequence 15323, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
;  APPLICANT: Deikman, Jill
;  APPLICANT:  Feng, Paul C.C.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Ziegler, Todd E.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(52274)B
;  CURRENT APPLICATION NUMBER: US/10/021,323
;  CURRENT FILING DATE:  2001-12-12
;  PRIOR APPLICATION NUMBER: US 60/255, 619
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 15323
;   LENGTH: 589
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3829-024-Q6-K6-G7
US-10-021-323-15323

  Query Match             22.5%;  Score 206.6;  DB 17;  Length 589;
  Best Local Similarity   65.0%;  Pred. No. 2.4e-43;
  Matches  370;  Conservative    0;  Mismatches  189;  Indels   10;  Gaps    4;

Qy        107 TTTTGGGTGCTTGGTTCAGTCCATTTGCCCTGAGGGTGCAGATTGCCCTTAACCTCAAGG 166

Db         24 TTTTGGGTTCCTGGCCGAGTCCATATGCGATGAGGCCAAGGATTGCCCTTAACCTCAAAT 83

Qy        167 GTTTGGATTATGAGGTTGTTGAAGAGACTTT---GAATCCCAAAAGTGAATTGCTTCTTA 223

Db         84 CCCTCACTTACGAATACATTGAAGAGAAGTTATGGGAAGGCAAAAGCGAGCTTCTTCTCC 143

Qy        224 AGTCCAACCCTGTGCACAAG-AAAATCCCAGTTTTCTTCCATGGAGAT---AAAGTCATA 279

Db        144 AATCAAACCCTGTTTACAAGTAAAATCCCAGTCCTCATTCACGGGGATAACAAGCCAATC 203

Qy        280 TGTGAATCTGCAATCATAGTTGAGTACATAGATGAGGTTTGGTCCAACAATGCTCTCTCC 339

Db        204 TGCGAATCTCTCATCATCGTACAATACATCGACGAGGCTTGGTCTTCC---GGTCCTTCT 260

Qy        340 ATCCTTCCACAAAATGCATATGATCGAGCTAATGCCCGATTTTGGGTTTCTTACATCGAT 399

Db        261 ATTCTTCCTTCTGATCCCTGTGAACGTGCTGTCGCTCGGTTTTGGGCTGCATATCTTGAC 320

Qy        400 GACAAGTGGCTTACGTCCTTGAAAAGTGTTCTAGCGACTGAAGATGATGAGGCAAAGAAG 459

Db        321 GACAAGTGGTTCCCAGCGGTGAAAAGTATAGGAGTCGCTAAAGGAGAGGAAGCAAAGAAA 380

Qy        460 CTACACTTTGAGCAAGCGGAAGAAGTGCTTGAGAAGGTGGAAGAAGTGTTCAACAAGTGC 519

Db        381 GCAGCAATAGCGCAAGTGGAAGCAGGGTTGGCTTTGATGGAAGAAGCATACGACAAGTGC 440

Qy        520 AGTGAAGGGAAGGCCTATTTCGGAGGAGATACGATTGGATTTGTTGACATTGGTTTTGGA 579

Db        441 AGCAAAGGGAAGGGTTACTTTGGTGGGGATGAAGTTGGGTACCTTGATATAGCTTTTGGG 500

Qy        580 AGCTTTTTGAGTTTCATTAGAGTCTCAGAGAATATGAATGAAAGAAAATTGCTTGATGAA 639

Db        501 AGCTTCTTGGGATGGCTTAGAGTGACGGAGAAGTTCAATGCGATGAAGCTGCTGGACGAG 560

Qy        640 ACGAAGTACCCTGGTTTGACCCTATGGGC 668

Db        561 GGGAAGACGCCTGGTCTGGTAAAATGGGC 589


Search completed: October 13, 2004, 07:14:09
Job time : 536 secs
```